Application cover sheet

Name of Respondent: PricewaterhouseCoopers Advisory Services LLC

Mailing Address: 101 Seaport Boulevard	**City/Town:** Boston	**State:** Massachusetts **Zip code:** 02210
Telephone: +1 (617) 530 5000	**Fax:** N/A	**Web Address:** https://www.pwc.com/us/en.html
Primary Contact for Clarification: Praveen Krishna		**Primary Contact E-mail Address:** Praveen.Krishna@pwc.com
Authorized Signatory: Ravi Kalakota		**Authorized Signatory E-mail Address:** Ravi.s.kalakota@pwc.com
Legal Status/Jurisdiction: LLP		**Respondents UEI No.** L5BYHKR1JJ68

Massachusetts Technology Collaborative

Authorized Respondent's Signature and Acceptance Form

The undersigned is a duly authorized representative of the Respondent listed below. The Respondent has read and understands the RFP requirements. The Respondent acknowledges
that all of the terms and conditions of the RFP are mandatory, and that Respondent's response
is compliant with such requirements.

The Respondent understands that, if selected by the Mass Tech Collaborative, the Respondent
and the Mass Tech Collaborative will execute an Agreement specifying the mutual requirements of participation.
Respondent agrees that the entire bid response will remain valid for sixty (60) days from receipt
by the Mass Tech Collaborative.

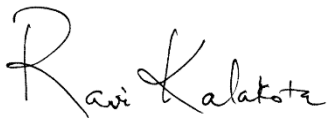
I certify that Respondent is in compliance with all corporate filing requirements and State tax
laws.

I certify that the statements made in this response to the RFP, including all attachments and
exhibits, are true and correct to the best of my knowledge.

I acknowledge and understand the procedures for handling confidential materials submitted to
the Mass Tech Collaborative as set forth in Section 3.1(d); I agree to be bound by those procedures and agree that the Mass Tech Collaborative shall not be liable under any circumstances for the disclosure of any materials submitted to the Mass Tech Collaborative
pursuant to this RFP or upon the Respondent's selection.

Respondent:

Notwithstanding any statements to the contrary in the MassTech AI hub Data Commons Collaborative Consulting Services RFP (the "RFP"), submission of this proposal is not an indication of PWC Advisory Services LLC] (PwC) willingness to perform the services under the terms of any existing master services agreement between the parties or to be bound by the terms presented in the RFP. This proposal in response to the MassTech's RFP does not constitute a contract to perform services and cannot be used to award a unilateral agreement. Final acceptance of any contract by PwC is contingent upon successful completion of PwC's acceptance procedures. Any engagement(s) arising out of this proposal will be subject to negotiation of a mutually satisfactory contract.

By: 

Name: Ravi Kalakota
Title: Advisory Partner
Date 10/07/2025

Notwithstanding any statements to the contrary in the MassTech AI hub Data Commons Collaborative Consulting Services RFP (the "RFP"), submission of this proposal is not an indication of PWC Advisory Services LLC] (PwC) willingness to perform the services under the terms of any existing master services agreement between the parties or to be bound by the terms presented in the RFP. This proposal in response to the MassTech's RFP does not constitute a contract to perform services and cannot be used to award a unilateral agreement. Final acceptance of any contract by PwC is contingent upon successful completion of PwC's acceptance procedures. Any engagement(s) arising out of this proposal will be subject to negotiation of a mutually satisfactory contract.

Application cover sheet

Name of Respondent: PricewaterhouseCoopers Advisory Services LLC

Mailing Address: 101 Seaport Boulevard	**City/Town: Boston**	**State:** Massachusetts **Zip code:** 02210
Telephone:+1 (617) 530 5000	**Fax:** N/A	**Web Address:** https://www.pwc.com/us/en.html
Primary Contact for Clarification: Praveen Krishna		**Primary Contact E-mail Address:** Praveen.Krishna@pwc.com
Authorized Signatory: Ravi Kalakota		**Authorized Signatory E-mail Address:** Ravi.s.kalakota@pwc.com
Legal Status/Jurisdiction: LLP		**Respondents UEI No.** L5BYHKR1JJ68

MassTech's AI Hub Data Commons Collaborative

October 7, 2025

October 7, 2025

PricewaterhouseCoopers Advisory Services LLC

101 Seaport Boulevard. Boston, Massachusetts 02210

www.pwc.com/us

Notwithstanding any statements to the contrary in the MassTech AI hub Data Commons Collaborative Consulting Services RFP (the "RFP"), submission of this proposal is not an indication of PWC Advisory Services LLC] (PwC) willingness to perform the services under the terms of any existing master services agreement between the parties or to be bound by the terms presented in the RFP. This proposal in response to the MassTech's RFP does not constitute a contract to perform services and cannot be used to award a unilateral agreement. Final acceptance of any contract by PwC is contingent upon successful completion of PwC's acceptance procedures. Any engagement(s) arising out of this proposal will be subject to negotiation of a mutually satisfactory contract.

A message from your PwC leaders

Thank you for the opportunity to submit our proposal to join MassTech's Data Commons Collaborative Design, Implementation, and Launch. Our experience deploying scalable platforms in a cross-sector ecosystem and our know-how in supporting conveners of such ecosystem in the construct of public-private partnership uniquely positions us to increase the likelihood of success for this initiative.

The following are some key points to consider as you read through our proposal:

- **Rapid assessment and delivery using AI with value realization:** You will notice a repeated reference to quick upfront assessments, accelerators leveraging AI technology for faster and smarter delivery. This is not intended to repeat the work that produced your DCC priorities, but rather to build upon it to verify points of view from participants and iterate in an organized manner deliver on value initiatives as well as identify opportunities for quick wins.
- **Critical controls and decision points:** In the program overview, you will note how we propose to use the Value Realization Office to help us manage program progress and risks. Our goal is to institute bi-monthly milestones and periodic communication with executive sponsors to maintain alignment throughout the entire organization on this critical set of initiatives.
- **A multidisciplinary team of experienced practitioners:** We have tapped experienced partners with ~20 years of experience in Healthcare, Research, Academics, and Finance to work on this program. While we are committed to sourcing a local Massachusetts team whenever possible, I am prioritizing expert national partners to lead the team to increase our likelihood of success.
- **Deep data and AI experience:** We bring cross-sector Data and AI expertise partnering with clients to design data strategies, implement scalable platforms, and unlock actionable insights. Our governance and ethics approach provides rigorous data governance, privacy protection, bias detection and mitigation, and transparent AI stewardship to deliver measurable outcomes.
- **Accelerating value capture in Year 1:** Our six-month timeline for MVP go-live accelerates value initiative demonstration in Year 1. We will earn the confidence of platform stakeholders by employing controlled introductions to establish a "safe harness" around initial users of the production environment to get feedback from early adopters, foster stakeholder trust, and avoid institutional fatigue.

As we undergo this journey together, we believe that the strategies employed above will help MassTech unleash the full potential of its data and AI Hub while gaining momentum needed for your sponsors to champion this initiative. Please do not hesitate to contact us with questions or feedback about our proposal. We look forward to discussing the next steps.

Sincerely,

Rima Safari
Advisory Partner

Oneil Rego
Advisory Partner

Ravi Kalakota
Advisory Partner

Table of Contents

1 Executive summary

2 Organizational background

3 Our perspective on DDC

4 Our approach

5 Our experience

6 Project team and fees

7 Our capabilities and accelerators

8 Appendices

1

Executive Summary

The Data Commons will transform Massachusetts' civic heritage into digital infrastructure for AI-driven innovation

	Massachusetts extends its tradition of shared civic spaces into the digital age.	There is an imperative to build a new "common" where data is stewarded for trust and fairness.	You are seeking a trusted advisor to assist with launching a pivotal civic space: the Data Commons.	With phased expansion, the Data Commons becomes a living engine of Massachusetts' innovation.	It will be built for sustainability, governance, and measurable impact so Massachusetts continues to lead.
Then	The Boston Common is America's first public park, open to all and stewarded for community benefit.	Early challenges of managing land fairly and preventing overuse.	The Common became a central gathering place for all citizens.	Boston Common remains vital centuries later as a symbol of shared stewardship.	Boston Common remains vital centuries later as a symbol of shared stewardship.
Now	The Data Commons Collaborative is a shared data platform convening datasets, talent, and innovation under fair governance.	Fragmented, siloed data limits equitable access and slows innovation.	Governed, equitable data infrastructure with fairness checks and shared access.	The Data Commons Collaborative is built to endure as an inclusive, trusted institution.	The Data Commons Collaborative is built to endure as an inclusive, trusted institution.

Our proposal at a glance

What we are proposing

DCC Marketplace
Engage users, refine vision, finalize designs, and build and deliver the Marketplace and integrations

DCC Platform
Establish a robust data platform and accelerate capabilities across prioritized sector and external datasets

DCC Enabling Services
Collaborate with the broader stakeholder team to establish governance, compliance, security, responsible AI, and operational capabilities to run DCC

Value Realization Office
Continuously and iteratively assess and formulate value initiatives that drives use case, build, and accelerated delivery

Our approach is tailored to MassTech

User focused & use-case led	Data personas enabled	Modular technology – Gen AI driven	Iterative – feedback oriented

Our solution sets the foundations for accelerated value delivery

2026 - 1st 180 days	2026 – Q3/Q4	2027 - Q1 and beyond
Foundations + Launch	Enrichment and growth	Scaling and accelerating consumption
• DCC Platform enabled with catalog, metadata, and synthetic data management services • DCC Marketplace launched for limited datasets and limited users with controlled introduction • Operational readiness, security, and privacy controls to enable launch • Robust product back log with prioritization • Integrated governance	• Onboarding additional use cases • Expansion of Data Platform • Additional features onboarded • Citizen-led development • Expand Marketplace • Enable Data Sciences • AI Workbench • Observability	• Continue to scale • Increased automation to drive business and data operations efficiencies • Accelerated adoption • Organization level standards • Compliance Management System • Ecosystem O

Scope and pricing commitment in this proposal

Outcomes we are focused on

User centric value delivery and launch

Enable Data Platform and Marketplace value delivery, anchored on prioritized data sets to enable AI innovation and trusted workspace for incentive-driven collaborations

Speed to delivery

Rapidly configure and deploy the mesh and marketplace environments with proven Gen-AI driven accelerators (e.g. PwC's Marketplace Accelerator) and state-of-the-art capabilities (PwC's Intelligent Data Modernizer accelerator)

Gain early stakeholder buy in to drive trust

Relentlessly focus on MassTech's success - integrated design sessions to enable early stakeholder buy in and establish champions to promote the data usage through the Marketplace

Scale governance across delivery, operational and data management

We thrive in multi-vendor environments and our integrated governance between data/product owners/stewards and partners will drive ownership and accountability

We envision a progressive journey where the Data Commons evolves from a secure catalog to a platform that powers the Massachusetts AI ecosystem

We recommend first focusing on delivering an MVP within the first six months to prove value, receive feedback, and avoid early technical debt

A solid, scalable foundation within the first year

- Stand up FAIR metadata, RBAC, audit trails, and governed APIs
- Release semantic search and dataset-level lineage
- Introduce baseline synthetic data capability
- Enable Workbench entry and session control

Augment foundation with intelligence within two years

- Add data quality scores, enhance usage dashboards, and integrate fairness or bias monitoring
- Expand federation connectors and pushdown patterns; keep hosting selective
- Establish controls over AI bias and fairness

Operate at scale with trusted AI by Year 3

- Turn on streaming and event-driven pipelines
- Offer developer SDKs and stable APIs so users build on the Commons, not around it
- Tie platform growth to workforce and startup programs so usage translates to outcomes

Day One principles					
	Federation first, host when it improves user value: Governed pushdown first	**Hybrid compute with MGHPCC: C**loud runs catalog and APIs; AICR runs GPU or HPC jobs	**AI services that interoperate:** Model gateway to managed endpoints or self-hosted	**Storage control and trust:** Placement and residency rules, versioning and SIEM integration	**Transparency on usage:** Event capture and Commons Credits ledger

2

Organizational background

PwC is perfectly positioned to be **the** experienced partner to design, build, launch, and scale the Data Commons Collaborative

Building trust in society and solving important problems

PwC is the largest professional services network in the world, offering Assurance, Tax, and Advisory services to our clients. We were formed in 1998 by the merger of two firms — Price Waterhouse and Coopers & Lybrand — each with historical roots going back some 170 years.

Our network comprises of approximately 370,000 professionals in 149 countries and 656 locations. We draw upon the skills and experience of 75,000 professionals and 4,000 partners in 79 offices in the United States. PwC network firms are privately owned by their partners and principals.

Your engagement team will bring the collective power of this network to find the right resources to address your specific needs. We'll bring you a global perspective with an in-depth understanding of local, state, and federal issues.

149
Countries

656
Locations

Our revenue in USD:
FY24: $55.4B
FY23: $53.1B
FY22: $50.3B
FY21: $45.1B
FY20: $43.0B

370,000+
Our people

We have the global reach, experience, and skill you need

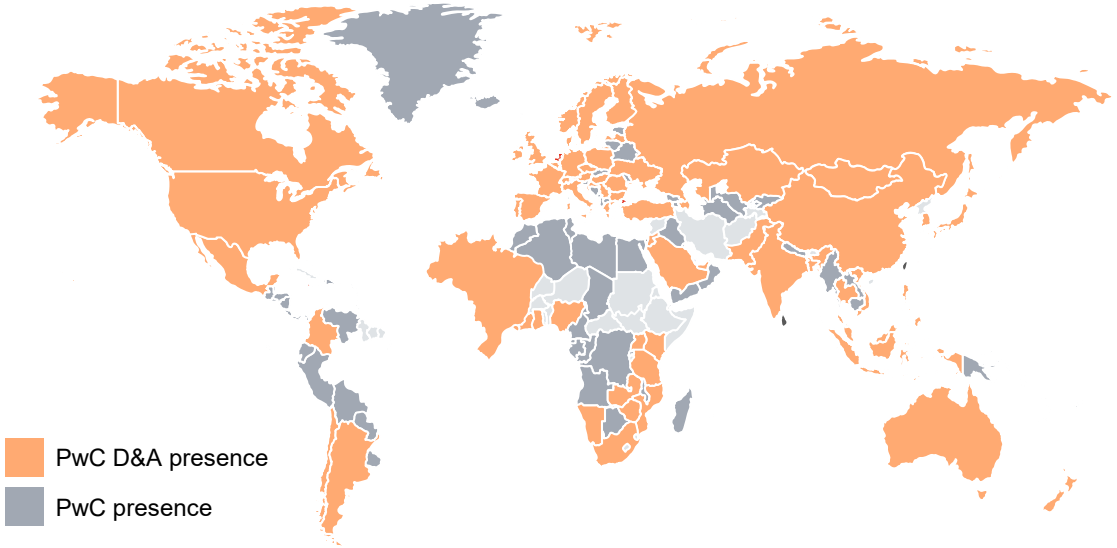
Our clients
86% of the Fortune Global 500 are our clients

PwC's Data, Analytics and AI (DA&AI) Practice Snapshot

Our team spans the globe, bringing the right talent at the right time to support MassTech's needs

35+ territories, 5 industries

- Financial Services (FS)
- Retail & Consumer
- Industrial Mfg. & Automotive (IM&A)
- MedTech, Healthcare & Pharma
- Technology, Media and Telecom (TMT)



Expertise
18,000+ D&A engagements in FY22
400+ Fortune 500 clients

Data Benchmarking
11,000+ metrics across **35,000+** companies
320+ data sets

Innovation
PwC Labs for AI & EmTech, Rapid PoC

Responsible AI
Customizable frameworks, tools to **reduce gaps between promise and performance of AI**

Accelerator Apps
2,000+ data analytic apps and accelerators for client engagements

Products
PwC Data Platform, Insights Platform, Workbench with **cloud-first approach**

Relationships
100+ ventures, **40+** technology major vendors and start-ups and **50+** universities

Experience
42 Experience and innovation centers spread across **every major continent**

Leader in FY21 Gartner Customer Choice & FY22 Forrester Data Management and AI

Collaboration with institutions, NGOs, industry bodies and governments

Leading-edge research and co-development of frameworks with universities

28,000+ FTEs in D&A and AI

8,200+ AI professionals

$6B+ Global D&A and AI revenue

$1.5B+ 2022 Global AI revenue

We leverage Global Acceleration Centers, augmented with our strategic Joint Business Relationship Partners, to deliver tailored, scaled, and efficient services

Global Acceleration Centers + **Joint Business Relationships** = **Services with Local Presence and Global Reach**

Global Acceleration Centers are hubs of innovation, capability, and efficiency strategically located around the globe and packed with skilled experts across every domain we serve.

JBR partners bring niche expertise, specialized skills, and local market presence that complement our delivery team. Part of delivery model for strategic clients include working with our clients to identify JBRs that provide local market collaboration. All JBRs are carefully vetted and onboarded to the PwC delivery model to drive delivery, quality, and a seamless client experience.

You will notice throughout this proposal our approach to using PwC accelerators to deliver on your needs. We will also collaborate with you to identity and onboard new JBR partners based on your priorities and strategic roadmap. We have identified the below key entities as the initial list to address DCC requirements:

aws	DATADOG	snowflake	Subsalt
Cloud Partner	Observability	Data Platform	Synthetic data

Along with our global presence, we take pride at putting our purpose in action at a local level to build trust in society and solve important local problems

Here is how we've invested in our local greater-Boston relationships over the past year.

Skills-based volunteering

$2.3M **Total pro bono value***

provided by our local people to organizations through fee-waived projects, board service, and **Skills for Society** projects/workshops

21.1k **Volunteer hours**

provided by our people to organizations within our Boston communities

198 **Nonprofit board seats**

held by our people in Boston communities

Inclusion Network participation

In Boston, our people are active across our **12 Inclusion Networks.**

- Abilities and Allies Inclusion Network
- Black Inclusion Network
- Environmental Inclusion Network
- Inter-Belief Inclusion Network
- Latino Inclusion Network
- Pan Asian Community Inclusion Network
- Parenting Network
- Secondees, Expatriates, and Immigrants Network
- Shine — Inclusion Network for LGBTQ+ professionals
- Veterans Affinity Network
- Women's Inclusion Network
- Women in Technology Inclusion Network

Boston community relationships

Some of the organizations we are collaborating with to help solve society's most critical problems:

- Bottom Line
- Camp Harbor View
- Cathedral High School
- Cathleen Stone Island Outward Bound
- Cristo Rey Boston
- Notre Dame Christo Rey High School
- Girls Inc.

Charitable giving

$1.5M **Charitable donations**

by our people in Boston

$1.1M **Reimagine grants***

Awarded through the PwC Foundation in Boston

***As of last PwC Foundation fiscal year ending 9/30/2024**

3

Our perspective on Data Commons Collaborative

Our Data Marketplace solution is built with hyper-focus on usability, and our delivery methodology will holistically address DCC needs to maximize adoption

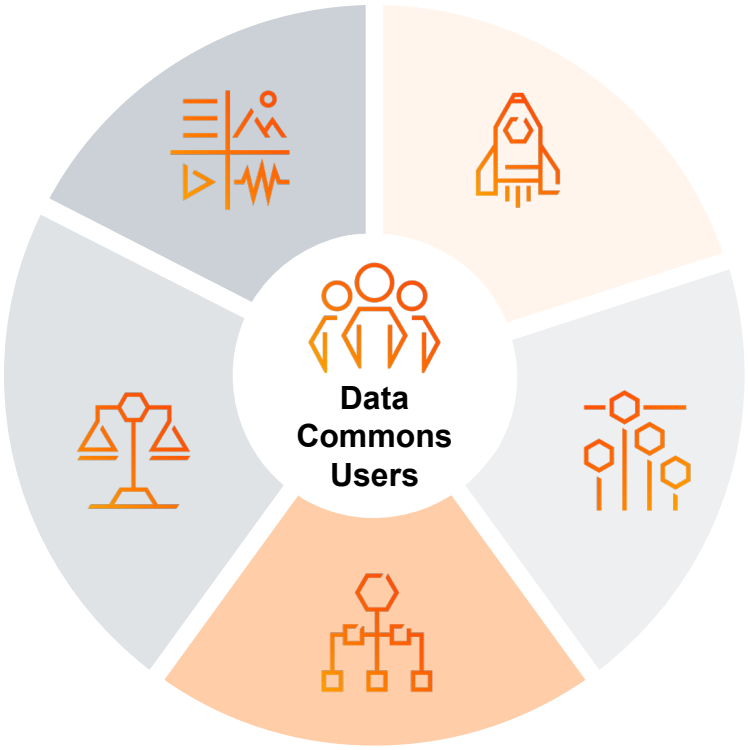
We instrument adoption and learn fast so we invest in what users actually use

We run discovery sprints with real users, prototype early, and close the loop with regular sprint demos.

We bake in governance and responsible AI as part of the product, not an afterthought

We deliver vertical slices that span UI, services, compute, and stores to prove value fast.

We confirm “host vs. federate” with experiments: Start federated, set performance, and adoption thresholds, promote to hosted only when users benefit.



Understanding the needs of the AI Hub’s Data Commons users is critical

We start with users because adoption depends on meeting real, varied needs, and not just standing up technology

A user-first design protects against a “field of dreams” outcome: success is measured by usage, not just by what we build

Early engagement and feedback loops enable the platform to evolve around practical value

By designing around user journeys, we make the Data Commons intuitive and indispensable from day one

Our proven Data Marketplace solution will help the Data Commons deliver an "Amazon-like user experience"

Find

Search available, categorized data assets

Preview and explore data samples

Request support to source/onboard new data

Access

Request access to "sensitive" data

Track progress on access approvals

Select and provision data products for use

Trust

Collaborate with "Experts" and other users

Understand relevance in the "context of use"

Rate and provide feedback on data products and features

Deliver

Define and publish "new and enhanced" data products

Accelerate analytical data preparation and automation

Facilitate data enrichment and quality management

Data Commons Data Lake

Users — the Data Commons should be designed and built to address a wide variety of data enablement needs, with user personas modeled on “data acumen and behaviors”

Casual Business User	Data Consumer and Requestor	Data Dabbler	Governance Upholder	Experienced Operator	Expert Superuser
I prefer high fidelity data representations and insights. I don’t have individual data needs that would require help from someone else.	I request help preparing or analyzing datasets from experienced or expert users. I know a lot about what I need, but I’m not working with data or coding myself.	I understand how to perform light data operations to fulfill my needs. I need help on more complex data operations from experienced or expert users.	I am a user that receives requests to provide others access to data. I also manage the governance and publishing of data, ensuring that data meets quality standards.	I know how to prepare and model data using advanced tools, and fulfil requests from others. I seek out experts for help on jobs when I need additional support.	I am an expert data operator who is well-versed in a variety of tools. I rarely need help, and support team members and data requestors.

Governance — it is likewise imperative for stewards to control risk in one place with policy-led access, continuous observability, fairness checks, and verifiable lineage

1

Federated sign-in, centralized policy

Data Commons users leverage existing agency/organization logins, while we control roles, consent, and risk in one place.

2

One audit trail across all IdPs

Show who accessed what, when, and why, with evidence ready for oversight.

3

Responsible AI practices and monitoring

Models and datasets carry cards and evaluations, so bias is detected proactively, and trust is visible to everyone.

4

Policy-led placement and residency

Data assets live where policy and value say it should, with sensitive copies kept near MGHPCC when needed.

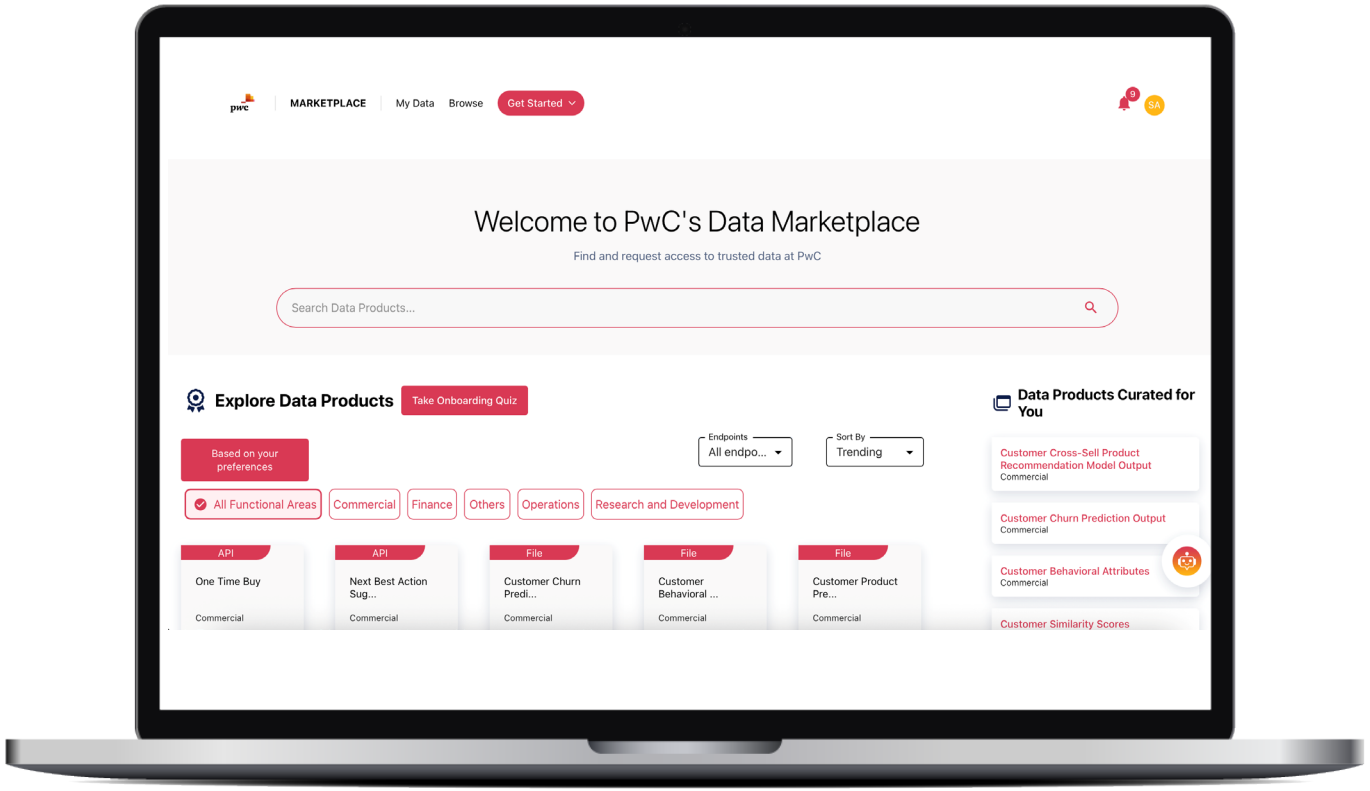
5

Prepared incident response

If issues arise, alerts and playbooks drive action and a report goes out rapidly with proof.

Data Commons Data Lake

We will help you incorporate a Data Marketplace as your DCC front door and white label it as DCC Marketplace



Diverse stakeholders can initiate data sharing and exploratory analysis in minutes (versus days/weeks) by quickly discovering available assets, products, and tools

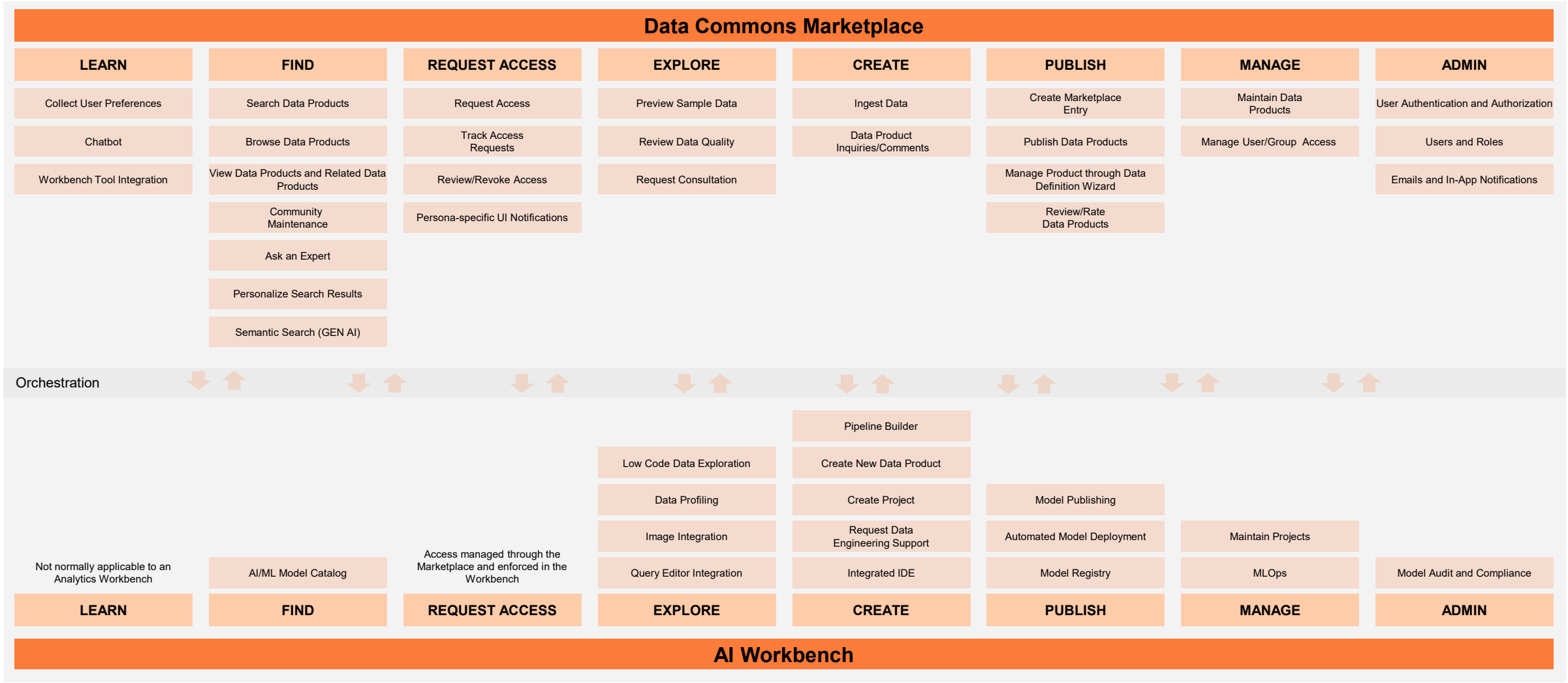
~60% acceleration of value from data-centric use cases by enabling the reuse of previously engineered products

Significant reduction in data access/exposure risks through centrally managed data classification, secure access patterns and usage monitoring

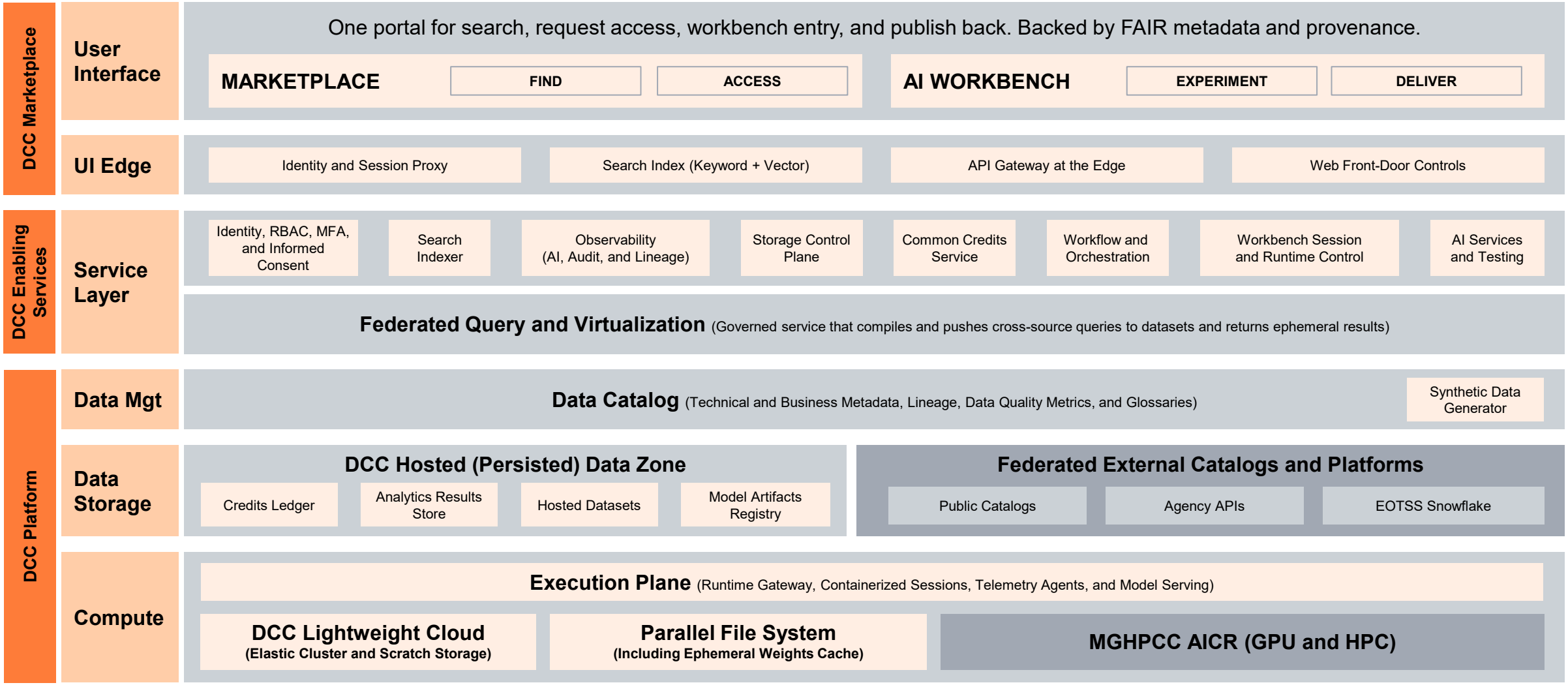
~30% reduction in total cost of ownership through consolidation redundant of data consumption patterns, repositories, tools, and support services

Sense of community and greater collaboration across all data practitioners (producers, stewards, and consumers) resulting in productivity gains

Data Marketplace features deep dive — through our experience, we have seen common capabilities move users from discovery to collaboration to publishing back, within a governed ecosystem



Your solution — to enable a seamless user journey, the frontend is backed by a governed, federation-first platform that directs compute to MGHPCC and hides the complexity



4

Our approach

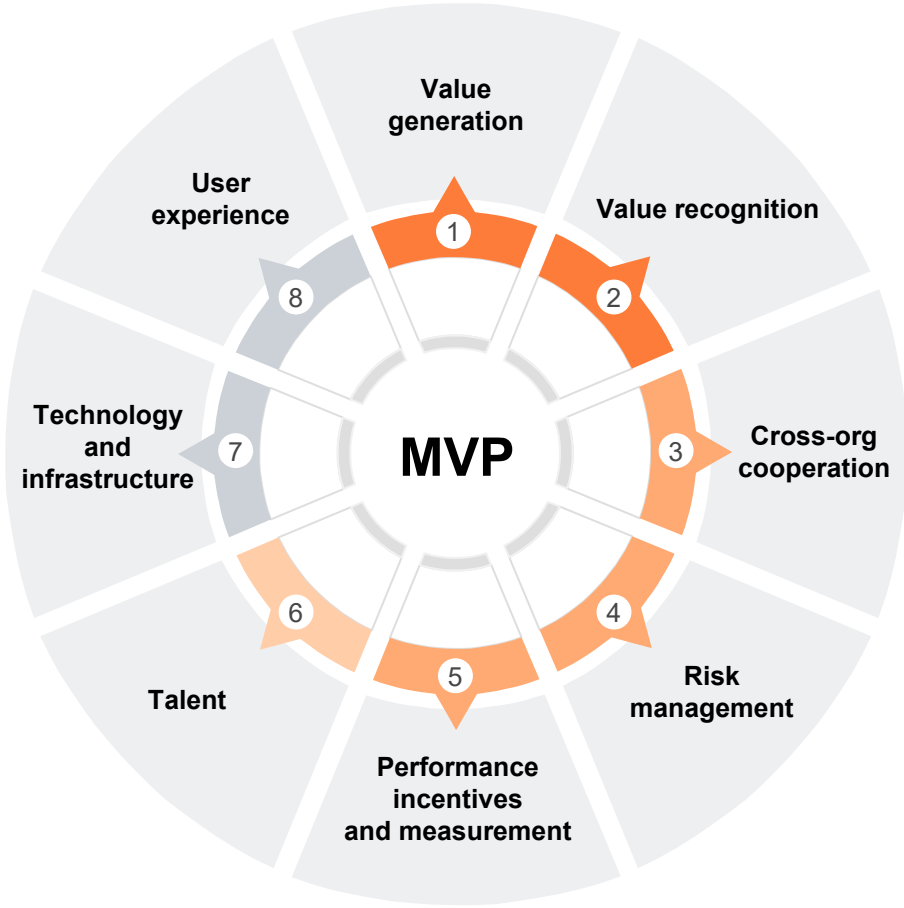
MassTech's vision for the AI Hub and DCC will require more than just technology, hence we believe an MVP (Minimum Viable Product) approach that emphasizes the below dimensions is critical for success

8 How will MassTech services be effectively **consumed** by each stakeholder?
(PwC developed apps with intuitive user experiences that were easy to use and engaging for users)

7 Do you have the right technical capabilities to deliver? How do the systems integrate and utilize technology to **achieve shared benefit**, and who pays?
(PwC deployed the platform, apps, and services, and navigated shared benefit across functions and geographies)

6 How will MassTech attract and **retain the superior talent in tech, AI, and analytics**?
(PwC focused on upskilling the organization and identifying data and AI skill gaps that needed to be enriched)

5 How will targeted **performance be motivated and measured?**
(PwC developed targeted incentives [tailored to varied stakeholders] and performance measures)



Massachusetts Technology Collaborative AI Hub

Value | Operations | People | Technology

1 How will the AI and Data Products **generate value rapidly and consistently** across public, industry, research, and education stakeholders?
(PwC designed "value initiatives" to make sure business value was attained)

2 How will the value be experienced across stakeholder groups to **motivate sustained adoption and funding**?
(PwC established an organizational change management team, and created a quarterly newsletter to disseminate value stories across the organization)

3 How will MassTech **mobilize a complex ecosystem** (e.g., colleges, clinics, hospitals, x-sector industry) to achieve state and organization goals?
(PwC convened a complex health ecosystem to collaborate to achieve improved patient outcomes)

4 How will "data" drive **high impact, responsible use of AI** to fulfill business model objectives?
(PwC helped the client establish a Data and Analytics Governance body to guide responsible use of data and insights)

We propose deploying a limited scope MVP in six months to demonstrate value, gain stakeholder confidence, and set the foundation for successful expansion

Key assumptions

Data: Limited internal datasets from priority sectors (life sciences, healthcare, robotics, and education) and Limited external datasets will be ingested, curated, and hosted

Target users: Pilot users will be limited to a reasonable group; representing each cohort (e.g. student, industry, academia)

Features: Focus will be on core data catalog, metadata, hosting, and front-end experiences, all remaining features will be on a prioritized backlog

Enabling Services: Scope of enabling services will be designed and rolled out in support core platform using minimum required principles

DCC Marketplace

An "Amazon-like" interface for users to find, access, and trust data.

- Marketplace build:
  - Foundations
  - Core find and learn features
  - Data product details
  - Admin interface
  - Access control workflow
  - Community features
- UI/UX customizations for MassTech
- Data Catalog integration
- Identity and access control integration for launch
  - User self-service registration
  - Integration with existing directory service

DCC Enabling Services

Common services to support the Marketplace and platform that includes data and AI governance, cybersecurity, privacy, and identity management.

- Design Integrated Data and AI Governance
  - Design and configuration, including bias and fairness controls
  - Gather requirements for AI review process system and observability platform
  - Design Data Access and Use Committee operating model
  - Draft acceptable use of data with AI policy
- Define AI Solution Security requirements
- Scope observability and monitoring requirements and design for future DataDog integration

DCC Platform

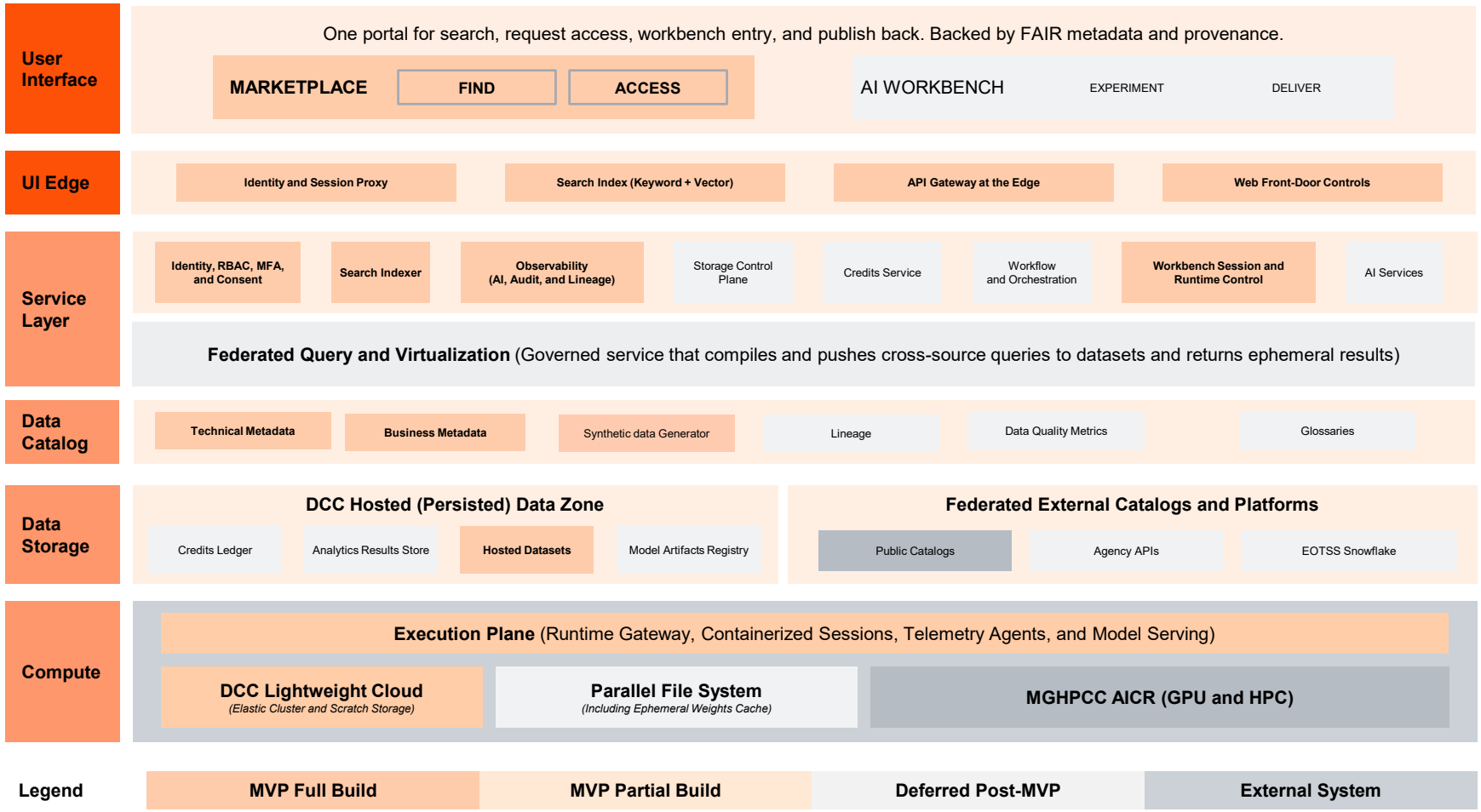
Scalable Data Platform to ingest, host, and curate datasets for the DCC marketplace.

- Cloud environment (Infrastructure as Code – Terraform scripts) setup
- Release environments dev/Test/QA/Prod
- Dataset curation criteria framework
- Ingestion pipelines and curation for in-scope dataset
- SubSalt Synthetic data generation pipelines
- Data Catalog and Metadata solutions and test with in-scope dataset
- Foundational core infrastructure, API, and data security capabilities required for launch

Technology Operations

DevSecOps	Operational Readiness	Compliance	Controlled Introduction	Quality

The MVP will bring the Data Commons to life by standing up the core, enabling users to find and access trusted data



MVP architectural scope:

- Users will be able to search and browse for data, request access, and download data for use
- The MVP will run in one secure cloud tenant, hosting the frontend Marketplace, open-source data catalog, and batch-first data lake for end-to-end workflows
- Advanced capabilities, including federated query, AI workbench, model registry, and credits common are designed in this phase but activated later as identity and policy mature
- Data catalog, metadata management and synthetic data generation solutions will be integrated
- A hybrid cloud architectural pattern is assumed between AICR and preferred cloud vendor

We've structured the work into four focused workstreams aligned with your requested services, built to continue delivering value beyond the MVP

PwC workstreams

1. **DCC Marketplace –** Drive end-user experience and design/development of DCC features, AI workbench, and Data Commons credits per strategic roadmap
2. **DCC Platform –** Drive development of a modern data platform that will ingest, curate, and host datasets that will be published on Data Marketplace for consumption; Address compute and storage needs
3. **DCC Enabling Services –** Translate policy and compliance requirements into people, process, and technology and build a secure environment for AI innovation; Enable Data and AI governance
4. **Value Realization Office –** Project risk/issue mitigation, value assessment, quality assurance, training, change management, launch, and marketing support

MassTech services solicited

- **Program management and coordination**
- **Technical architecture and platform development**
- **Dataset curation and integration**
- **Synthetic data capability**
- **AI Fairness and Bias framework**
- **Security, governance, and compliance**
- **Testing, rollout, and adoption**

We will organize ourselves into a fit-for-purpose staffing model at each stage of the MVP build; value realization office will have oversight across activities and outcomes

Establish a Technical Design Authority — cross-functional architects from MassTech + PwC (for MVP and beyond)

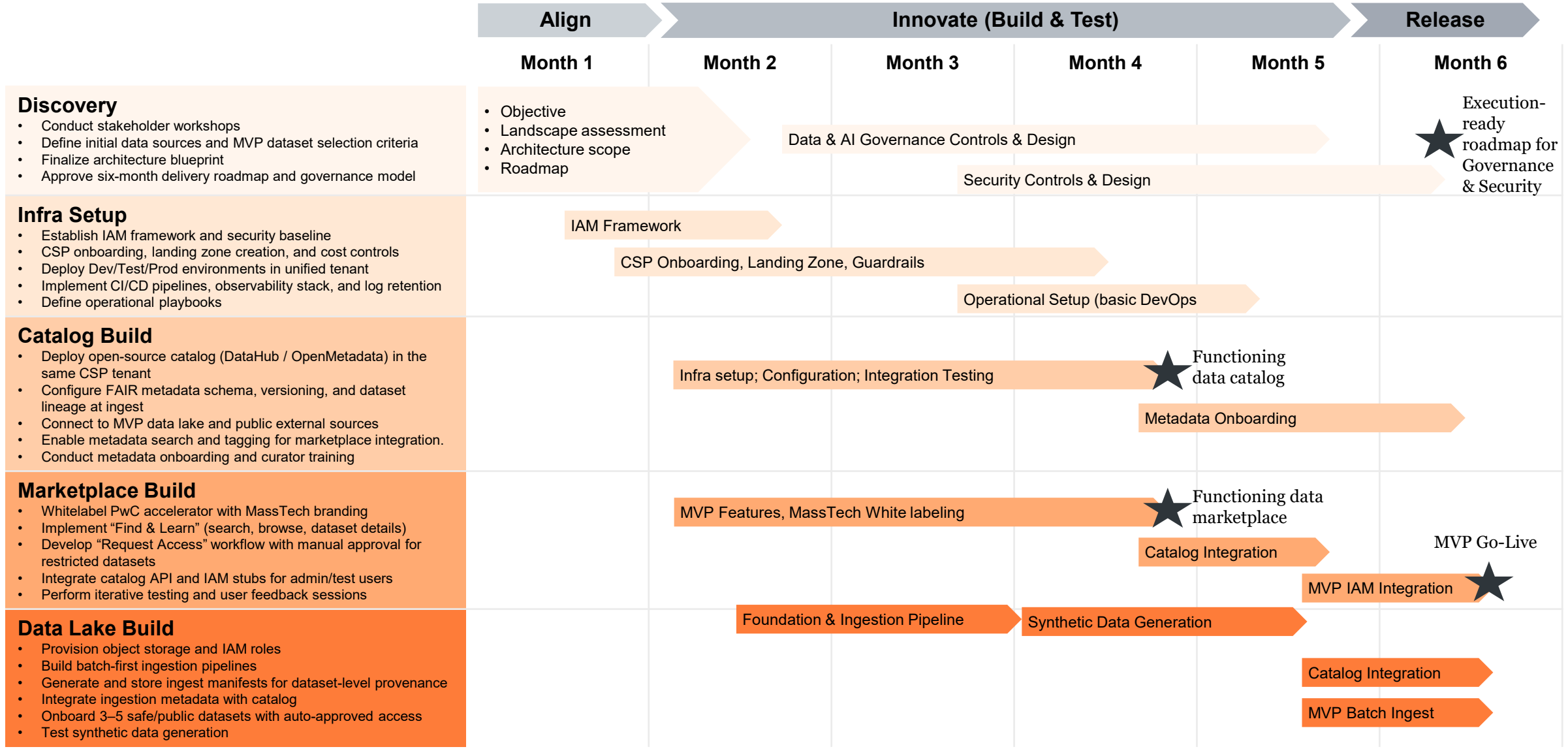
Activate PwC's propriety agentic AI SDLC framework and tooling to accelerate build

MVP go-live — activate managed services/operations

	Mobilize	Design	Implement	Launch
DCC Marketplace	• Establish program governance structure, including a steering committee • Activate project management and value realization office • Establish necessary contracts • Onboard/Procure necessary technology vendors and resources • Engage stakeholders and establish feedback cadence • Establish DevSecOps tooling and pipeline	• Web Application design • Rapid front-end prototyping using Figma/Framer • User and admin UI/UX • Identity and access design	• Establish Agile ceremonies • Two-week sprint cadence • Activate cross-functional build pods • Develop web front end • Develop and deploy Data Marketplace MVP v1.0 • Data platform build scope follows data being onboarded for MVP	• Controlled Introduction of MVP [Pilot] — tech validation + data validation + intended user in intended environment validation • Iterative monthly release planning for modules/ capabilities beyond MVP • Training and documentation • Public launch support • Operational KPIs and dashboard
DCC Platform		• Low-level architecture design/Data flow design • Data catalog/MDM tool design • Synthetic data generator design • Cloud components/services		
DCC Enabling Services		• Core security design (data + cloud) • Monitoring and observability tool design • Compliance framework • AI Fairness and Bias frameworks	• Deploy foundational data, API, and infrastructure security • Integrate RBAC and IAM components	

Value realization office

MVP timeline: DCC Marketplace and DCC Platform detailed Technical workplan


Align
Innovate (Build & Test)
Release
Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Discovery
• Conduct stakeholder workshops
• Define initial data sources and MVP dataset selection criteria
• Finalize architecture blueprint
• Approve six-month delivery roadmap and governance model
• Objective
• Landscape assessment
• Architecture scope
• Roadmap
Data & AI Governance Controls & Design
Security Controls & Design
Execution-ready roadmap for Governance & Security
Infra Setup
• Establish IAM framework and security baseline
• CSP onboarding, landing zone creation, and cost controls
• Deploy Dev/Test/Prod environments in unified tenant
• Implement CI/CD pipelines, observability stack, and log retention
• Define operational playbooks
IAM Framework
CSP Onboarding, Landing Zone, Guardrails
Operational Setup (basic DevOps
Catalog Build
• Deploy open-source catalog (DataHub / OpenMetadata) in the same CSP tenant
• Configure FAIR metadata schema, versioning, and dataset lineage at ingest
• Connect to MVP data lake and public external sources
• Enable metadata search and tagging for marketplace integration.
• Conduct metadata onboarding and curator training
Infra setup; Configuration; Integration Testing
Functioning data catalog
Metadata Onboarding
Marketplace Build
• Whitelabel PwC accelerator with MassTech branding
• Implement “Find & Learn” (search, browse, dataset details)
• Develop “Request Access” workflow with manual approval for restricted datasets
• Integrate catalog API and IAM stubs for admin/test users
• Perform iterative testing and user feedback sessions
MVP Features, MassTech White labeling
Functioning data marketplace
Catalog Integration
MVP Go-Live
MVP IAM Integration
Data Lake Build
• Provision object storage and IAM roles
• Build batch-first ingestion pipelines
• Generate and store ingest manifests for dataset-level provenance
• Integrate ingestion metadata with catalog
• Onboard 3–5 safe/public datasets with auto-approved access
• Test synthetic data generation
Foundation & Ingestion Pipeline
Synthetic Data Generation
Catalog Integration
MVP Batch Ingest

While MVP will provide the foundation, we will use our capabilities and accelerators to address DCC requirements holistically

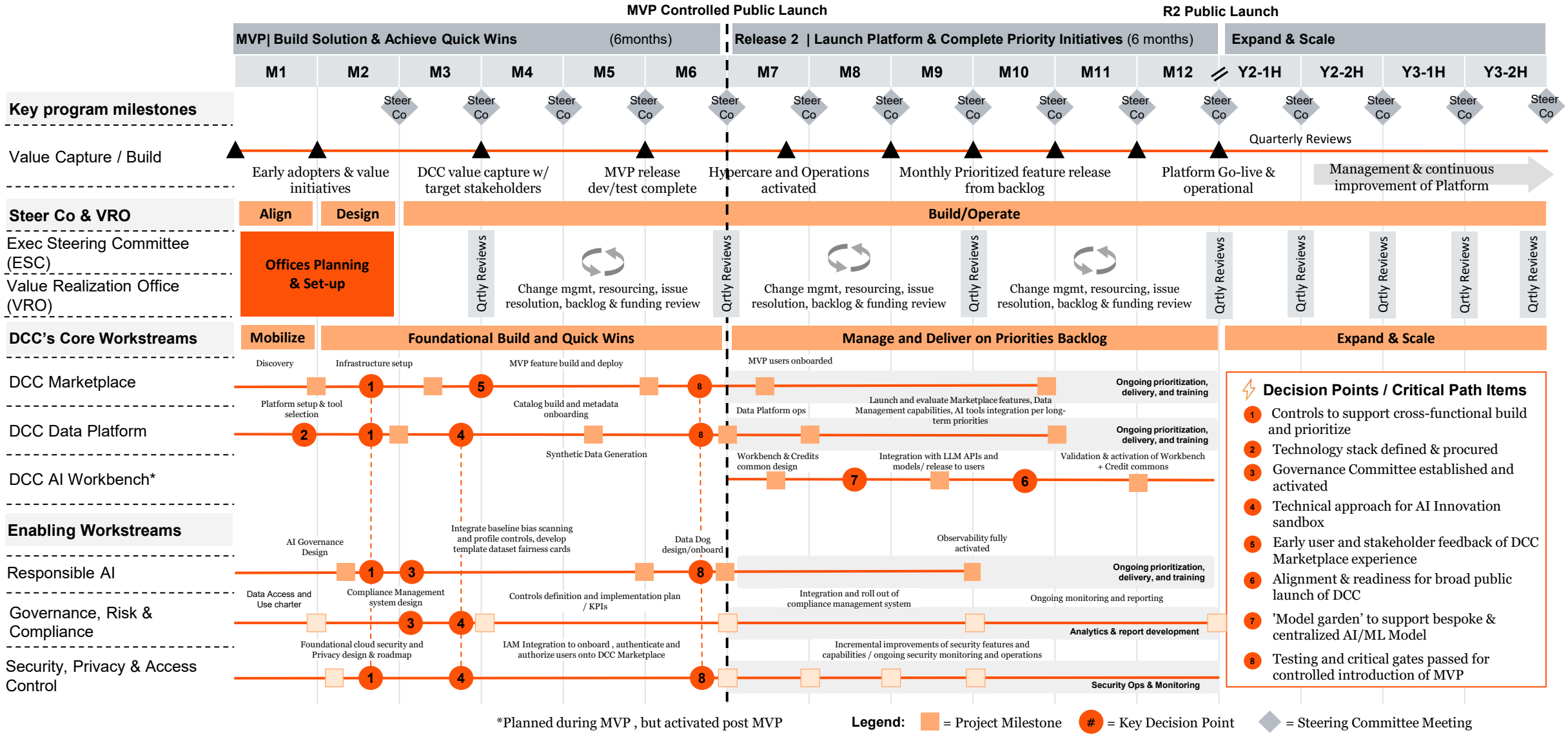
DCC requirements	Our proposed solution	Select PwC Accelerators
External data sets and data capabilities **Website functionality for hosted data and access**	Using PwC's proven Data Marketplace deployment playbook and codebase, we will build a white-labeled DCC Marketplace solution with an iterative feature release plan beyond MVP. MassTech can customize the available features and add custom ones as needed to the backlog. This solution will cater to various user personas and provide a data and AI workbench where projects can be spun up on demand, with outputs generated and released. Our preference is to build this on AWS; however, the solution is cloud-agnostic and can be deployed on both Azure and GCP infrastructures.	• **Data Marketplace** ("Amazon –like user experience" to help data stakeholders find, access, use and collaborate on common datasets in a trusted environment)
Common credits system	We propose to build a bespoke module integrated with the data marketplace to address these requirements. Our AI-driven application development framework can rapidly prototype, iterate, and build this solution.	• **AI-enabled SDLC toolkit** (deliver accelerator in code development and deployment)
Data catalog and metadata management system **Data curation plan**	We will leverage MassTech's current investments in Snowflake (with strengths including interoperability, access control and governance, and data previews/profiling) and integrate it with CKAN and/or Dataverse to address metadata management (descriptive, technical, and metadata standards). We will evaluate Data Marketplaces' publishing, versioning, and related capabilities to be selective and strategic about incremental solutions. Our Intelligent Data Modernizer uses AI-driven profiling and assessments of datasets prior to ingestion and processing.	• **Intelligent Data Modernizer** (AI driven data platform development and deployment toolkit)
Synthetic data generation	We are working with Subsalt Inc. to address your needs around synthetic data generation. The AI-driven tool will assist in generating not only statistically similar copies of real datasets (e.g., healthcare AI use cases) but also fit-for-purpose synthetic data to train new AI models and simulate digital twins of processes and systems.	• Homegrown custom synthetic data generation tools in collaboration with SubSalt Inc.

Refer to section 7 for details around our solution, capabilities and accelerators to meet DCC requirements

While MVP will provide the foundation, we'll use our capabilities and accelerators to address DCC requirements holistically

DCC requirements	Our proposed solution	Select PwC Accelerators
AI fairness and bias capability **Governance policies and ethical guidelines**	We will apply PwC's Responsible AI Framework and Bias Testing Toolkit to define fairness dimensions, bias detection methods, and ethical data-use principles that establish trust in AI systems from design through deployment. We will embed observability and bias monitoring controls at the data and model levels to ensure transparent and explainable AI outcomes. Leveraging PwC's AI governance and risk management accelerators, we will develop actionable policies, transparency mechanisms, and accountability structures aligned to regulatory expectations and responsible AI best practices. We will define requirements, design, and roadmap for an observability solution (i.e., our JBR with Datadog will allow for a free trial period in subsequent phases, if selected by Mass Tech).	• PwC Responsible AI Framework • PwC AI Risk & Model Cards • PwC AI Bias, Model Testing, & Evaluation Toolkit • PwC Draft Policies & RACI Templates • JBRs with Observability Platforms (e.g., Datadog)
Governance requirements	We will establish a DCC governance framework aligned with state, federal, and emerging AI regulations, defining clear roles, decision rights, and oversight through data stewards, fairness reviewers, and an AI/Data governance council. The framework will specify data-sharing models and access tiers—with federated, restricted, controlled, and open-access options—and outline provider and user responsibilities and terms of use. We will implement consent management (tokenized, revocable, time-bound) and acceptable-use controls that operationalize nondiscrimination and fairness principles. Policy will be translated into technical controls including RBAC/MFA, lineage, immutable audit logs, encryption, data retention and transparency reporting. Finally, we will establish monitoring dashboards, governance KPIs, and training programs to strengthen trust and community capabilities.	• PwC Compliance Management System (CMS) Framework • Template policies and SOPs for operations • PwC Sample AI Roles and Responsibilities and RACIs • PwC Transparency & Assessment Toolkit • PwC Quality Framework
Security and access controls	We will use our templates and playbooks to integrate the data marketplace solution with MassTech's SSO solution, with capabilities like self-service registration, password reset, Federation and multi-factor authentication. We have pre-built integrations for RBAC and dataset permissions within the solution, which we will enable and tailor to MassTech's requirements. Our cybersecurity engineers and practitioners follow zero-trust architecture guidelines for cloud infrastructure and data security and will implement the necessary controls in accordance with industry leading practices and MassTech's security policy.	• PwC Zero Trust Cloud Security Reference Architecture • PwC Privacy Framework • PwC IAM Integration libraries and playbooks

Our roadmap and milestones for DCC's broader objectives and initiative

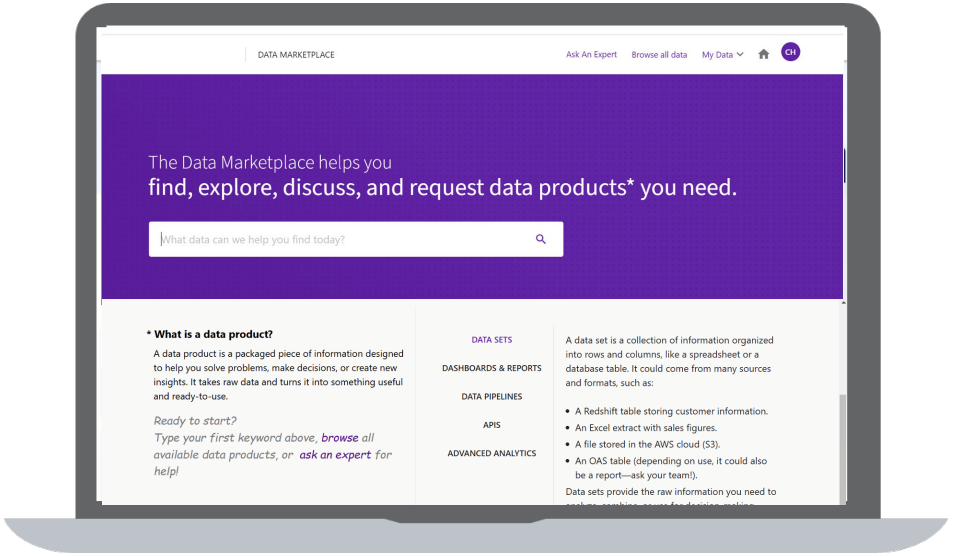

MVP Controlled Public Launch
R2 Public Launch
MVP| Build Solution & Achieve Quick Wins (6months)
Release 2 | Launch Platform & Complete Priority Initiatives (6 months)
Expand & Scale
M1 M2 M3 M4 M5 M6 M7 M8 M9 M10 M11 M12 Y2-1H Y2-2H Y3-1H Y3-2H
Key program milestones
Steer Co
Value Capture / Build
Quarterly Reviews
Early adopters & value initiatives
DCC value capture w/ target stakeholders
MVP release dev/test complete
Hypercare and Operations activated
Monthly Prioritized feature release from backlog
Platform Go-live & operational
Management & continuous improvement of Platform
Steer Co & VRO
Align
Design
Build/Operate
Exec Steering Committee (ESC)
Value Realization Office (VRO)
Offices Planning & Set-up
Qrtly Reviews
Change mgmt, resourcing, issue resolution, backlog & funding review
DCC's Core Workstreams
Mobilize
Foundational Build and Quick Wins
Manage and Deliver on Priorities Backlog
Expand & Scale
DCC Marketplace
Discovery
Infrastructure setup
MVP feature build and deploy
MVP users onboarded
Ongoing prioritization, delivery, and training
Platform setup & tool selection
Catalog build and metadata onboarding
Launch and evaluate Marketplace features, Data Management capabilities, AI tools integration per long-term priorities
DCC Data Platform
Data Platform ops
Synthetic Data Generation
DCC AI Workbench*
Workbench & Credits common design
Integration with LLM APIs and models/ release to users
Validation & activation of Workbench + Credit commons
Enabling Workstreams
AI Governance Design
Integrate baseline bias scanning and profile controls, develop template dataset fairness cards
Data Dog design/onboard
Observability fully activated
Responsible AI
Data Access and Use charter
Compliance Management system design
Controls definition and implementation plan / KPIs
Integration and roll out of compliance management system
Ongoing monitoring and reporting
Governance, Risk & Compliance
Analytics & report development
Foundational cloud security and Privacy design & roadmap
IAM Integration to onboard , authenticate and authorize users onto DCC Marketplace
Incremental improvements of security features and capabilities / ongoing security monitoring and operations
Security, Privacy & Access Control
Security Ops & Monitoring
Decision Points / Critical Path Items
1 Controls to support cross-functional build and prioritize
2 Technology stack defined & procured
3 Governance Committee established and activated
4 Technical approach for AI Innovation sandbox
5 Early user and stakeholder feedback of DCC Marketplace experience
6 Alignment & readiness for broad public launch of DCC
7 'Model garden' to support bespoke & centralized AI/ML Model
8 Testing and critical gates passed for controlled introduction of MVP
*Planned during MVP , but activated post MVP
Legend: = Project Milestone # = Key Decision Point = Steering Committee Meeting

5

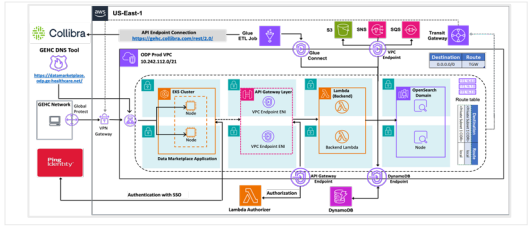
Our experience

A global MedTech organization launched an enterprise Data Marketplace to enable AI at scale, thereby unlocking cost synergies, efficiency, and revenue gains

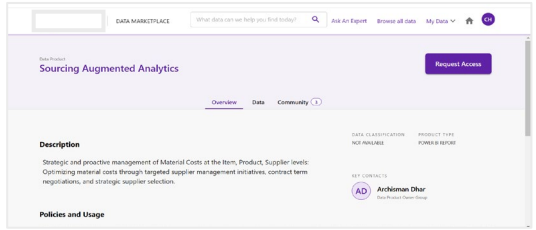
- A leading life sciences company was investing in an enterprise-scale data modernization effort to evolve its cloud data and analytics capabilities but lacked a unified experience for discovering, accessing, and managing trusted data products across the enterprise.
- Leadership prioritized data democratization, reuse, and governance as strategic enablers for analytics, AI, and operational efficiency.


DATA MARKETPLACE
Ask An Expert
Browse all data
My Data
The Data Marketplace helps you
find, explore, discuss, and request data products* you need.
What data can we help you find today?
* What is a data product?
A data product is a packaged piece of information designed to help you solve problems, make decisions, or create new insights. It takes raw data and turns it into something useful and ready-to-use.
Ready to start?
Type your first keyword above, browse all available data products, or ask an expert for help!
DATA SETS
DASHBOARDS & REPORTS
DATA PIPELINES
APIS
ADVANCED ANALYTICS
A data set is a collection of information organized into rows and columns, like a spreadsheet or a database table. It could come from many sources and formats, such as:
A Redshift table storing customer information.
An Excel extract with sales figures.
A file stored in the AWS cloud (S3).
An OAS table (depending on use, it could also be a report—ask your team!).
Data sets provide the raw information you need to

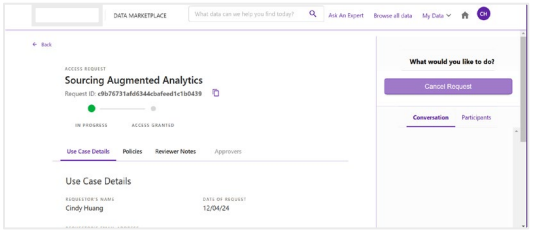
Solution



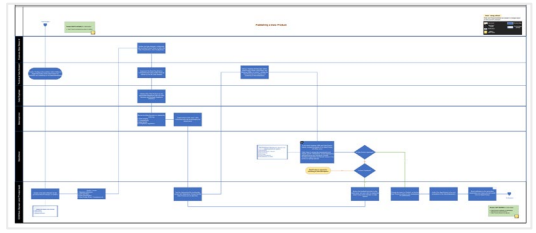
Launched a centralized marketplace for enterprise data products, **accelerating discovery, reuse, and decision making**


DATA MARKETPLACE
Ask An Expert
Browse all data
My Data
Sourcing Augmented Analytics
Request Access
Overview
Data
Community
Description
Strategic and proactive management of Material Costs at the Item, Product, Supplier levels: Optimizing material costs through targeted supplier management initiatives, contract term negotiations, and strategic supplier selection.
Archisman Dhar
Policies and Usage

Enabled data use by exposing related products, lineage, and usage details across business units



Provisioned automated access workflows and role-based controls to **reduce manual handoffs and delays**



Defined and **activated the data product life cycle** by operationalizing how products are created, reviewed, and published in the Marketplace

Results:

The Data Marketplace laid the foundation for faster, more scalable delivery of analytics and AI by enabling:

- **Fewer hours spent** searching for and validating data
- **Increased reuse of data and models,** reducing duplicate efforts and metric inconsistencies
- **Improved visibility and accountability,** with clear lineage, ownership, and access

CASE STUDY

Addressing the rising prevalence of diabetes through an expanded, community-based care delivery model that achieved improved disease control in patients receiving remote monitoring

Ecosystem Convener for Data Partnership

The challenge

The University of Texas System (UT System) sponsored a Diabetes and Obesity Control (DOC) program to address the growing prevalence of diabetes in South Texas. The UT System developed and deployed a program called Real-world Education, Detection, and Intervention (REDI) to demonstrate the beneficial impact of providing high-quality, convenient, and affordable chronic health care services to patients using PwC's product, the DoubleJump Interchange™ (DJI).

Why it matters:

Consumer retail data and purchasing habits are increasingly linked to wellness, prevention, and patient outcomes.

Client quote (publication)

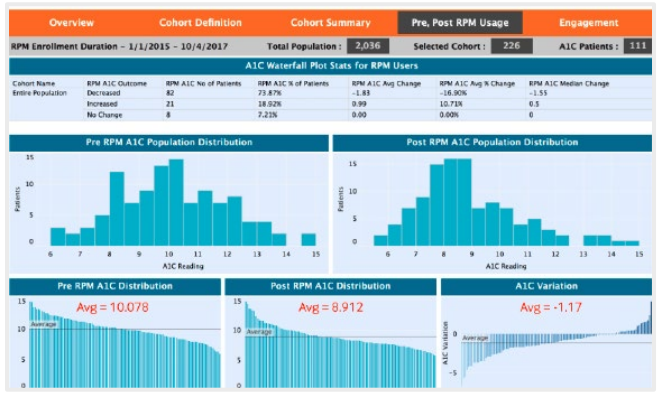
L Chin, JB McCormick, RS Greenberg. Convening a Digitally Enabled Ecosystem to Address the Chronic Disease Burden of an Underserved Community. New England Journal of Medicine - Catalyst (2018)

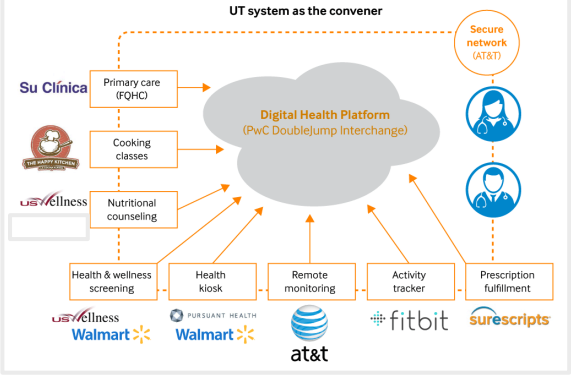
Impact and results

- 8,248 patients were screened for diabetes during weekly, four-hour free screenings at 13 retail stores throughout South Texas.
- 7,786 finger-stick tests were performed; 21% showed elevated readings (>100 mg/dL fasting; 199 mg/dL non-fasting).
- 349 HbA1c tests were performed; 41% of patients with diabetes (HbA1c > 6.5) were referred to primary care for final diagnosis.
- For patients on remote monitoring and diagnosed with diabetes and hypertension or another comorbidity (n = 203), we observed:
  - A 1 percentage point decrease in average HbA1c — in patients with type 2 diabetes, this represents an estimated ~$1,000 in annual cost savings in the US
  - A 14-point decrease in systolic BP — reducing the number of heart attacks by 21%, strokes by 37%, and all deaths from cardiovascular disease by 25%.

Delivering a connected health ecosystem for chronic disease control to reduced HbA1c and blood pressure, especially in those patients with remote monitoring

PwC's healthcare data cloud platform (DJI) was used to convene a diverse ecosystem of traditional and new health economy patient data (e.g., EMR, SDoH). Analytics of these data generated insights that transformed care delivery workflows, enhanced patient activation, and produced statistically significant clinical improvements associated with reduced costs.





The delivered solution

The REDI model applied established services and technologies (see diagram) to enable an expanded care team to provide chronic care to patients outside traditional care facilities. The DoubleJump Interchange™ integrated the resulting treatment data into a longitudinal patient record that could be accessed by the primary care provider to make treatment decisions, which were then carried out by the local care team. PwC served as the neutral broker during contract negotiations and configured a seamless workflow across the expanded care team.

CASE STUDY

A data-driven and AI-enabled approach to design an unprecedented well-care system focused on improving individual health span

AI-enabled bespoke Model of Care to improve health span of individuals

The challenge

Today's sick care system is overrun by expensive electronic health record (EHR) systems that are constrained by incomplete diagnosis and treatment data and are unable to inform optimized, individualized treatment plans tailored to regional requirements.

Neom's vision was to build a well-care system supported by a robust health profile that includes individual genomic, clinical, social determinants of health, and behavioral data. Neom's patient-first, person-centered philosophy empowers residents with access to their health data and decision support services, enabling them to make informed, personalized decisions with their clinicians about treatment and lifestyle options, while also contributing to advancing research in extending healthy life years.

Client quote

"Applying AI in healthcare can dramatically improve clinician experience and patient outcomes"

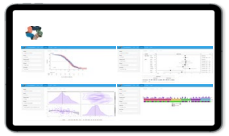
Impact and results

- Building a bespoke design using gamification and AI for healthcare, taking into account regional preferences and simplifying the UI/UX experience for end users, while helping physicians practice at the top of their license.
- A model of care based on wellness and supported by a command & control center to continuously monitor and intervene, when necessary, to maintain the health span of NEOM citizens.
- A consortium of academic institutions, technology companies, innovation hubs, hospitals, and clinicians from across the US and Saudi Arabia coming together to address global healthcare inequities.

Individual Co-Pilot

Clinician Co-Pilot

Clinician Data Workbench



Individualized Population Management

Digital Twin

Highly dimensional digital representation of a person

Proprietary Foundation Model

Pre-trained EHR foundation model, fine-tuned with supplemental contextual data

AI-powered decision support

Dynamic linking of P-360 with clinical guidelines to inform treatment options

The delivered solution

PwC worked with a consortium of academia, technology, clinical specialists, and innovation hubs to produce a technical design that included: a) an AI companion to motivate compliance and provide lifestyle guidance; b) an AI co-pilot to offer clinicians real time, AI-powered decision support at the point of care; c) tools to empower researchers with easy access to a low-code/no-code data insights workbench; and d) capabilities to enable early detection and intervention of time-sensitive health risks.

CASE STUDY

AI governance transparency and choice strategy for integration into a software company's products

Embedding responsible AI transparency, choice, and governance into a software company's AI-powered product ecosystem.

The challenge

PwC had the opportunity to collaborate with a multinational software corporation on their transparency and choice strategy. The company has made significant investments in AI and plans to embed AI capabilities across its products. These AI-powered solutions are built on vast amounts of data, including customer data. For the client, building customer trust through responsible AI practices, transparency, and choice was of utmost importance. The company wanted to explore how their peers approached transparency and choice and how these companies communicate the use of data and AI in their products, as well as the level of choice provided to users. The client was interested not only in identifying peers leading in this space but also in learning about the practices that have been most successful in building trust with customers and enhancing company brands.

Client quote (publication)

"Trust is paramount in the adoption of AI. And trust starts with transparency … we are introducing AI Transparency Cards … to clearly explain how each AI feature is built, the data used, and the benefits it offers."

Approach

PwC developed a robust transparency and choice assessment framework to evaluate how companies demonstrate transparency in their data and AI practices, as well as the level of choice offered to users. This framework was developed by identifying leading market practices and collaborating closely with the client to incorporate relevant criteria to inform decision making around data and AI use within their products. The process involved curating a diverse portfolio of companies and products, identifying key user types for assessment, and defining key dimensions to establish robust framework for evaluating transparency and choice practices at both the product and company levels.

After conducting evaluations of the products and companies within the peer cohort, PwC generated a detailed report outlining their findings. This report included a quantitative scoring system developed from a structured rubric, along with qualitative case studies summarizing key practices and distinctive product features. PwC also provided thorough analyses for each company and product, including a review of the client's practices relative to their peers.

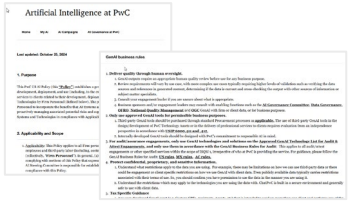
Lastly, leveraging these findings, the team developed recommendations for the client's transparency and choice strategy. These recommendations highlighted essential practices for building customer trust and identified opportunities for the client to differentiate themselves as a leader in responsible and trustworthy AI.

PwC Accelerators

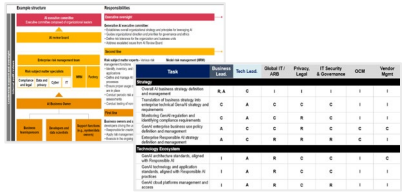
PwC's AI Risk Cards define a comprehensive risk taxonomy that highlight the potential to prioritize when enhancing AI capabilities
1. Model Risks
2. Data Risks
3. Infrastructure Risks
4. Use Risks

PwC AI Risk & Model Cards

PwC AI Bias, Model Testing, & AI Evaluation Toolkit



PwC Draft Template Policies



PwC Sample AI Roles and Responsibilities and RACIs

Impact

The PwC team delivered recommendations to the client, guiding them in implementing transparency practices that clearly communicate how AI is used within their products and highlight the resulting benefits for users. The recommendations also aimed to empower users by providing them with the necessary information to make informed choices about how their data is used within these products. By conducting this assessment, the client gained valuable insights into the performance of their peers regarding transparency and choice. They were able to embed common and notable practices that effectively build trust in the use of data and AI. The resulting report and controls proved instrumental in informing decision making at the C-suite level, shaping the client's data and AI strategy, and solidifying their position as a trusted AI partner in the market, thereby strengthening their brand.

References

Please refer to the attachment: Reference Form MassTech_PwC

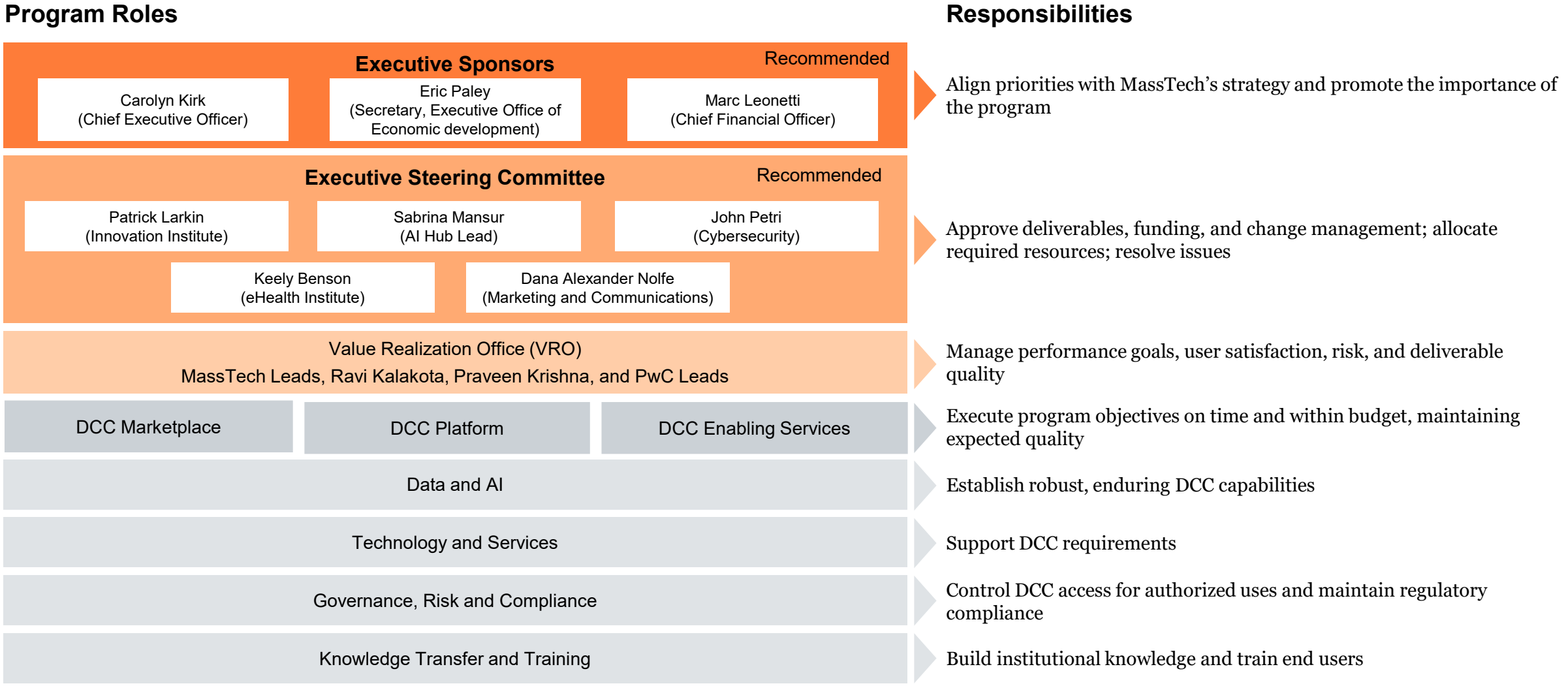
Engaging References

We have provided three client references (per your request) with whom we have closely collaborated to design and implement data strategies, data platforms, analytics, and AI in the healthcare sector, research, operations, and academia, as well as associated processes, to enable their organizations to achieve their performance and patient-impact goals. Out of respect for our clients' time and confidentiality, it is our policy to facilitate reference checks. When MassTech is ready to engage these references, please reach out to Ravi Kalakota (ravi.s.kalakota@pwc.com) and Oneil Rego (oneil.rego@pwc.com) to facilitate those discussions.

6

Project team and fees

Program Governance structure: Organizational commitment and engagement are required to meet expectations on time and within budget



We have hand selected the proposed team with relevant cross functional experience to help MassTech realize their DCC vision

Executive Leadership

MassTech Executive Leadership

- Ravi Kalokota
- Illana Blumenfeld

Executive Leadership provide strategic direction and endorses case for change or value proposition

Project Leads

MassTech Program Sponsor
AI Hub / Data Commons Leadership

- Praveen Krishna — Engagement Delivery Lead
- Martin Coffman — Responsible AI Lead
- Amol Chaudhari — Security & Privacy Lead

PwC SMAs

- Sachin Bajpai
- Karthik Umakanth
- Senthilkumar Jayachandran

Project leads, long with Subject Matter Advisors provide expertise and tactical direction for workstreams

Core Delivery Team

- Technical Architects
- System Owners
- Sector Leads
- Data Stewards

MassTech Project Manager

- Functional Stakeholders
- AI Hub enablement team
- Infrastructure Enablement Team

Sean Allen — Marketplace Delivery Lead

- Cloud Application Engineers
- UI/UX React Developers
- Data Catalog Developers
- Portal Delivery Lead

Spoorthi Purumala — Data Delivery & Platform Lead

- Platform Architect
- ETL Lead
- Data Engineers
- Data Modelers
- Testers
- Cloud Infra Engineers

Aamir Polli — Cross-Workstream Scrum Master

- Documentation Specialist
- Technical and Functional Testers
- Operations readiness Lead
- Compliance specialists
- Security Analysts
- Change Management specialist
- Communication specialist
- Training analyst

Core Delivery Team will comprise of all workstream leads and team members who will execute on scope of the functional components and contribute to successful release and adoption

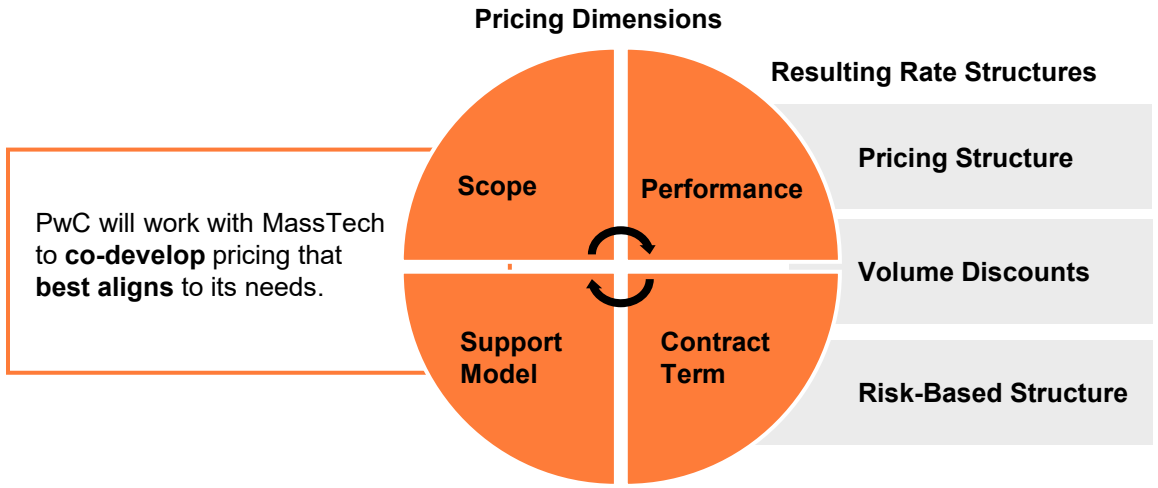
MassTech

PwC

Our competitive, modern pricing approach creates accountability and meets your transformation objectives

Our pricing approach for managed services blends traditional pricing methods with modern pricing structures to provide MassTech with a flexible arrangement that enables co-accountability for outcomes and cost efficiency.

Our commercial pricing proposal will be based on key dimensions that balance scope, performance, support model, and contract term. We have found that maintaining this balance is critical to designing a highly competitive pricing proposal aligned with your needs.



The Fixed Fee Advantage: Predictable, Aligned, and Outcome-Driven

- **Cost Certainty:** Provides full transparency and predictability by establishing the investment up front.
- **Outcome Focused:** Shifts emphasis from hours expended to the quality and timeliness of deliverables.
- **Aligned Interests:** Ensures our success is measured by results that matter to you.

.

Pricing Structure Examples

Rate Structure	Description	Best For	Example
Fixed Fee	Fixed monthly fee	Predictable services	• Service management and specialist services
Resource Unit	Cost per widget	User-/Device-/App-based support	• Price per ticket • Price per device • Price per application
Tiered Pricing	Multiple service tiers	Clients with environments with varying needs	• Gold/Silver/Bronze applications
Usage-Based	Based on actual consumption	Variable usage	• Discretionary pool for enhancements
Rate Card	Predictable hourly rate for specific skill types	Short-term project work	• Specific project with a defined deadline

Proposed fee structure for MVP

While many firms can provide platform architecture and technical support, PwC differentiates itself by delivering an **integrated approach optimized for long-term value creation**.

1. Tax and Incentive Optimization

- **R&D Tax Credits**: PwC's tax professionals can identify and maximize eligibility for federal and Massachusetts R&D credits, ensuring that project investments in AI development and synthetic data generation return tangible financial benefits.
- **Data Licensing & Transfer Pricing Guidance**: As data is aggregated from multiple sources (academic, private sector), PwC helps structure licensing and sharing agreements to minimize tax leakage, avoid cross-border disputes, and align with transfer pricing best practices.
- **Public Funding Alignment**: PwC advises on federal and state-level AI and innovation incentives, helping MassTech and ecosystem partners align Data Commons investments with broader grant, incentive, and compliance requirements.
- **State and Local Tax Efficiency**: With significant software, cloud, and consulting spend expected, PwC can provide state and local tax structuring to avoid unnecessary costs while ensuring compliance.

2. Responsible AI and Governance Leadership

- PwC brings a **proven Responsible AI framework** that goes beyond tool implementation (AIF360, Fairlearn) to include governance design, ethical AI oversight, and regulatory readiness.
- We provide **independent assurance services**—including audits of synthetic data quality, fairness dashboards, and data provenance—that create transparency and build trust with regulators, funders, and end-users.

3. Cybersecurity, Privacy, and Compliance at Scale

- Our cyber and privacy teams ensure the Data Commons meets the highest standards of **security (RBAC, MFA, encryption, audit trails)** while aligning with **GDPR, CCPA, and U.S. state privacy laws**.
- PwC brings **deep compliance expertise** in public–private data collaborations, helping MassTech avoid governance pitfalls as the Commons scales.

4. Ecosystem Engagement and Adoption

- PwC leverages its **Innovation Hubs, startup programs, and alliances with AWS, Microsoft, and Google** to accelerate industry adoption of the Data Commons.
- We offer structured **change management and stakeholder engagement**, ensuring that researchers, startups, government, and industry participants not only adopt the platform but also actively contribute to its success.

5. Full Life-Cycle Partnership

From **design → build → tax and funding optimization → responsible AI compliance → adoption and scale**, PwC offers unmatched breadth compared to competitors. This full-spectrum approach ensures that the Massachusetts AI Hub's Data Commons Collaborative will not only launch successfully but thrive as a national model for equitable and innovative AI development.

Fees

We want to work with you to give you exactly what you're looking for.

#	Work Product	MVP Fees (6 months)	Overall Fees (1-2 yrs)
1	DCC Marketplace	$874,087	$5M–$7M (estimated additional fees for design and implementation of the overall solution based on MVP outcome)
2	DDC Platform	$601,537	
3	DDC Enabling Services	$550,620	
4	**Total MVP Fees**	**$2,026,244**	
5	Total hours (range)	14,351–15,861	
6	Blenden RPH (range)	$127–$141	

Our fee for the scope of services requested is based on a fixed fee. Having reviewed your request, we believe that this pricing structure provides the suitable flexibility — and limits your responsibility for every hour worked. We're taking on the risk so that you have the benefit of paying for results rather than effort. We acknowledge that this structure may be different from our competitors, but we have confidence in our ability to deliver within the defined scope and schedule.

We want to work with you to give you exactly what you're looking for. Therefore, we've provided a set of options below that reflect your RFP request. Our next step would be to discuss these options with you to maintain alignment from Day One.

Pricing assumptions

- Effort estimated in the above MVP fee is to accelerate value delivery and does not include any additional costs related to technology vendor licenses or cloud consumption costs.
- We expect some travel expenses to be incurred to conduct mobilization/design sessions with MassTech stakeholders. We will work with MassTech to get such budgets pre-approved and be under a mutual agreed upon ceiling.
- Subcontractors and direct costs are not applicable per our approach and not included in the total MVP cost.
- The above estimate covers up-to 3 weeks of hypercare post MVP go-live, but doesn't cover any run or operate cost

7

Our capabilities and accelerators to address DCC requirements

Addressing DCC Requirements | **External data sets and data capabilities**

Our Proposed Solution/Assumptions

- Support both hosted data sets and federated discovery of external sources, with clear provenance and links to the authoritative owner
- Target auto-harvesting from CKAN, Dataverse, Socrata, and ArcGIS Hub, plus manual curation
- Provide a data set directory with browse, metadata search, provenance, and detail pages aligned to a FAIR-oriented profile
- Use an API gateway with rate limits, quotas, and API keys; support bulk and REST- or SQL-like access patterns
- Prefer hosting high-value shared data sets in the DCC with large, sensitive, or fast-changing data sets remaining external but governed

MVP Scope of Services

- Inventory and prioritize external data candidates
- Select prioritized federated data sources for MVP launch
- Stand up harvesters for prioritized federated data sources, with schedules, de-duplication, and review queues
- Configure API gateway for external integration, quotas, keys, and monitoring
- Pilot federated access patterns to allow models or users to work with external data without copying it
- Establish data provider onboarding kit — licensing options, refresh SLAs, quality checks, and governance checkpoints

PwC Tools and Accelerators

- **PwC Data Marketplace** ("Amazon –like user experience" to help data stakeholders find, access, use and collaborate on common datasets in a trusted environment)

Addressing DCC Requirements | **Website functionality for hosted data and access**

Our Proposed Solution/Assumptions

- React portal on CDN/edge, CSP service layer, cloud-agnostic (AWS-recommended), OIDC/SSO-ready
- Search and find features — search/browse, data asset pages with metadata, provenance, preview, and link related data assets
- RBAC, consent, tiered request -> access workflow
- Create/publish new data assets within the Data Commons Marketplace
- User community with comments, expert consultations, data asset ratings, and feedback

MVP Scope of Services

- Find — keyword + GenAI-powered semantic search, browse/filters, user personalization
- Data asset details — metadata, policies, sanitized sample data, ownership contacts
- Request and access — submit/track access, review/revoke, in-app and email notifications
  Publish/Manage — GenAI-powered metadata capture for new data assets, admin console for
  usage analytics
- Deliver — Portal with above features and function
- Service layer — identity, search indexer, approval workflows, AI Workbench integration

PwC Tools and Accelerators

- **PwC Data Marketplace** ("Amazon –like user experience" to help data stakeholders find, access, use and collaborate on common datasets ina trusted environment)
- **PwC Canvas /PwC Code Intelligence** - SDLC tooling leveraging Gen AI for rapid prototyping and code development

Addressing DCC Requirements | **Commons credit system**

Our Proposed Solution/Assumptions

- Establish a credit-based, identity-bound system to allocate DCC data, APIs, and compute
- Credits map to resource units — API calls, GPU hours, storage, and egress
- Tiered roles with RBAC govern default quotas and capabilities
- Time-bound credits replenish on a fixed cadence to prevent hoarding
- All usage is metered, logged, and auditable with monthly transparency reports
- Incentives are aligned with equity and value goals, not only cost control

MVP Scope of Services

- Policy and design — define units, tiers, rate card, renewal rules, and exception process
- Metering and ledger — central ledger that normalizes usage across APIs, jobs, and storage
- Audit and reporting — immutable logs, monthly usage statements, and internal equity dashboard
- Admin workflows — credit issuance, transfer, revocation, appeals, and misuse holds
- Pilot execution — run allocations for limited cohorts (e.g., university labs, state agencies, start-ups)
- Roadmap — Design phase two for public dashboards, marketplace integration, and auto-replenishment

PwC Tools and Accelerators

PwC Canvas /PwC Code Intelligence - SDLC tooling leveraging Gen AI for rapid prototyping and code development

Addressing DCC Requirements | **Data catalog and metadata management system**

Our Proposed Solution/Assumptions

- Support implementation of a metadata-driven catalog to support Findable, Accessible, Interoperable, Reusable (FAIR) principles
- Organize, describe, and make hosted datasets discoverable with descriptive, technical, and provenance metadata
- Enable dataset publishing, discovery, and access with governance built in
- Adopt standards (Dublin Core, Schema.org, DCAT) to ensure interoperability
- Integrate with AI Compute Resources (MGHPCC AICR) and external systems through APIs

MVP Scope of Services

- Stand up metadata catalog platform and configure metadata schema (descriptive, technical, provenance)
- Help implement dataset ingestion workflows with publishing, review, and versioning controls
- Provide RESTful APIs and SDKs for programmatic access and integration with external systems
- Establish role-based access tiers (public, registered, restricted), with MFA and audit logging
- Deploy data set previews and basic profiling to support usability
- Onboard a curated set of initial data sets to validate platform functionality

PwC Tools and Accelerators

- **Intelligent Data Modernizer** - AI driven data platform development and deployment toolkit that has a set of 10 accelerators leveraging Generative AI , playbooks, code bases to rapidly ingest data and build data pipelines

Building MassTech's DCC data catalog through integrated metadata and lineage management

Illustrative

1 **Onboard** to DCC data mart (Silver zone)

Business metadata
- Product details
- Domain mapping

Technical metadata
- Source name and description
- Data profiling output
- Physical data structure

Operational metadata
- Volume
- Frequency
- Pipeline status

Lineage
- Domain mapping

2 **Cleanse and prepare** data (Gold zone)

Business metadata
- Data definitions
- Business rules
- Data integration rules
- Data quality rules
- Data governance roles
- Data security rules
- Naming standards

Technical metadata
- Technical data model
- Data architecture
- Data flows
- Data integration and mapping rules
- Standards

Operational metadata
- Process name

Lineage
- Column/Table transformations
- Source to target mapping

3 **Provide access and consume** data

Business metadata
- Data asset name
- Description
- Type
- Purpose
- Primary use case
- Source system list
- Business owner
- Data domain
- Business function
- Product hierarchy
- Access groups
- Associated reports

Technical metadata
- Technical rules
- Access control

Operational metadata
- Data profile
- Update/Refresh date
- Pipeline performance

Lineage
- Lineage visualizations

4 **Share and monitor** usage

Business metadata
- Associated regulations
- Privacy considerations/flags
- Compliance stakeholders/reference

Technical metadata
- Access information
- Sharing information
- Audit reports

Operational metadata
- Workflows for usage
- Data retention policies
- Schedule information
- Load statistics
- Usage statistics

Lineage
- Validation status changes

Addressing DCC Requirements | **Data curation plan**

Our Proposed Solution/Assumptions

- Establish clear inclusion criteria — high-value for AI/ML, widely used, open/restricted license, anonymized, or synthetic-ready
- Curate initial datasets across key domains — general purpose, NLP, health/life sciences, multimodal/speech, robotics, social/urban, and Massachusetts-specific data
- Leverage domain specialists to validate dataset quality, ethical use, and licensing compliance
- Balance global benchmark datasets (e.g., ImageNet, Common Crawl, MIMIC), with Massachusetts-specific public datasets

MVP Scope of Services

- Define and publish dataset inclusion framework and governance process.
- Identify, catalog, and document priority datasets aligned to FAIR principles.
- Validate licensing and data-sharing terms for open/restricted use.
- Onboard initial high-value datasets into the Data Commons platform.
- Enable metadata tagging (domain, sensitivity, provenance) for discoverability.
- Develop roadmap for phased expansion into domain-specific and synthetic datasets.

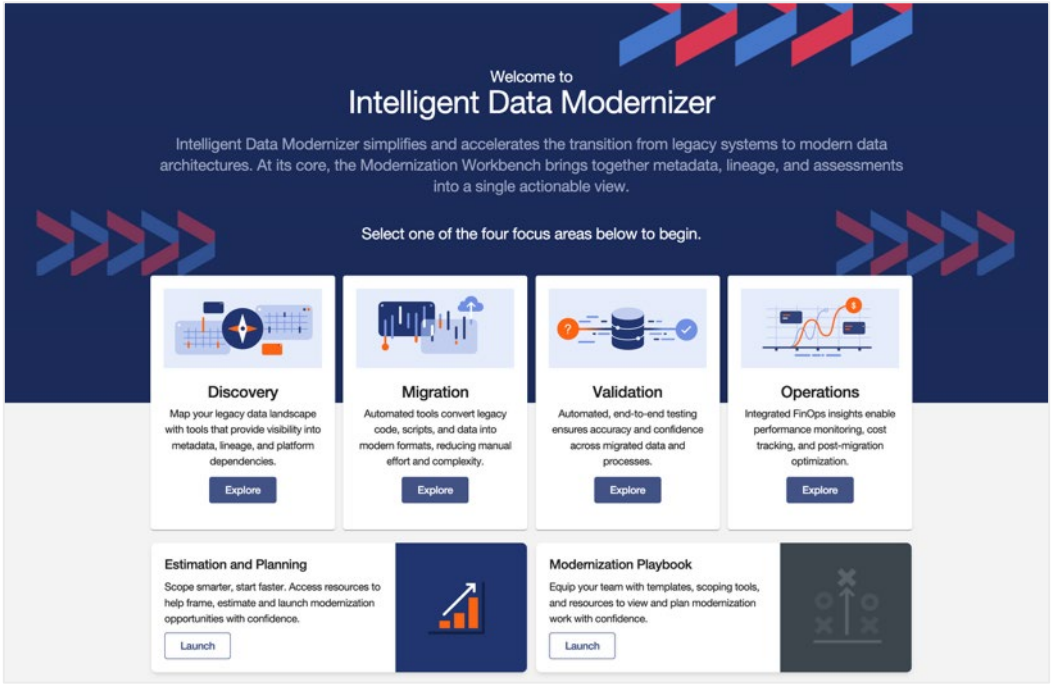
PwC Tools and Accelerators

- Anywhere to Cloud Discovery (Metadata & Lineage): Discovery and lineage collection from the legacy platforms
- Automated E2E Validation: Metadata driven framework, executes the testing process for data migration and ETL
- FinOps Manager: Monitor cost consumption for modern data platforms

We will use PwC's proprietary **Intelligent Data Modernizer** to accelerate your data ingestion, data curation, and data management — powered by agentic AI

The **Intelligent Data Modernizer** is a thorough workbench designed to streamline the modernization of legacy data systems. By leveraging advanced automation, AI-driven intelligence and agents, it dramatically **reduces the time, cost, and complexity of data discovery and migration processes.**

Achieve **3x faster** discovery and assessment, cutting months of effort down to weeks.


Welcome to
Intelligent Data Modernizer
Intelligent Data Modernizer simplifies and accelerates the transition from legacy systems to modern data architectures. At its core, the Modernization Workbench brings together metadata, lineage, and assessments into a single actionable view.
Select one of the four focus areas below to begin.
Discovery
Map your legacy data landscape with tools that provide visibility into metadata, lineage, and platform dependencies.
Explore
Migration
Automated tools convert legacy code, scripts, and data into modern formats, reducing manual effort and complexity.
Explore
Validation
Automated, end-to-end testing ensures accuracy and confidence across migrated data and processes.
Explore
Operations
Integrated FinOps insights enable performance monitoring, cost tracking, and post-migration optimization.
Explore
Estimation and Planning
Scope smarter, start faster. Access resources to help frame, estimate and launch modernization opportunities with confidence.
Launch
Modernization Playbook
Equip your team with templates, scoping tools, and resources to view and plan modernization work with confidence.
Launch

Key Capabilities and Value

Rapid time to value

Shorten Discovery and Assessment phases from months to weeks, allowing faster project execution and quicker realization of modernization benefits

Cost efficiency

Achieve over 30% reduction in delivery costs by identifying and eliminating redundant logic and unnecessary complexities, enhancing resource allocation

AI-powered insights

Leverage AI-driven profiling and validation to enhance decision making with real-time lineage and visual graph insights, improving accuracy and speed of strategic decisions

Seamless integration

Integrate seamlessly with existing DevOps workflows and orchestration tools, such as Airflow and Azure Data Factory, enhancing operational efficiency and reducing deployment friction

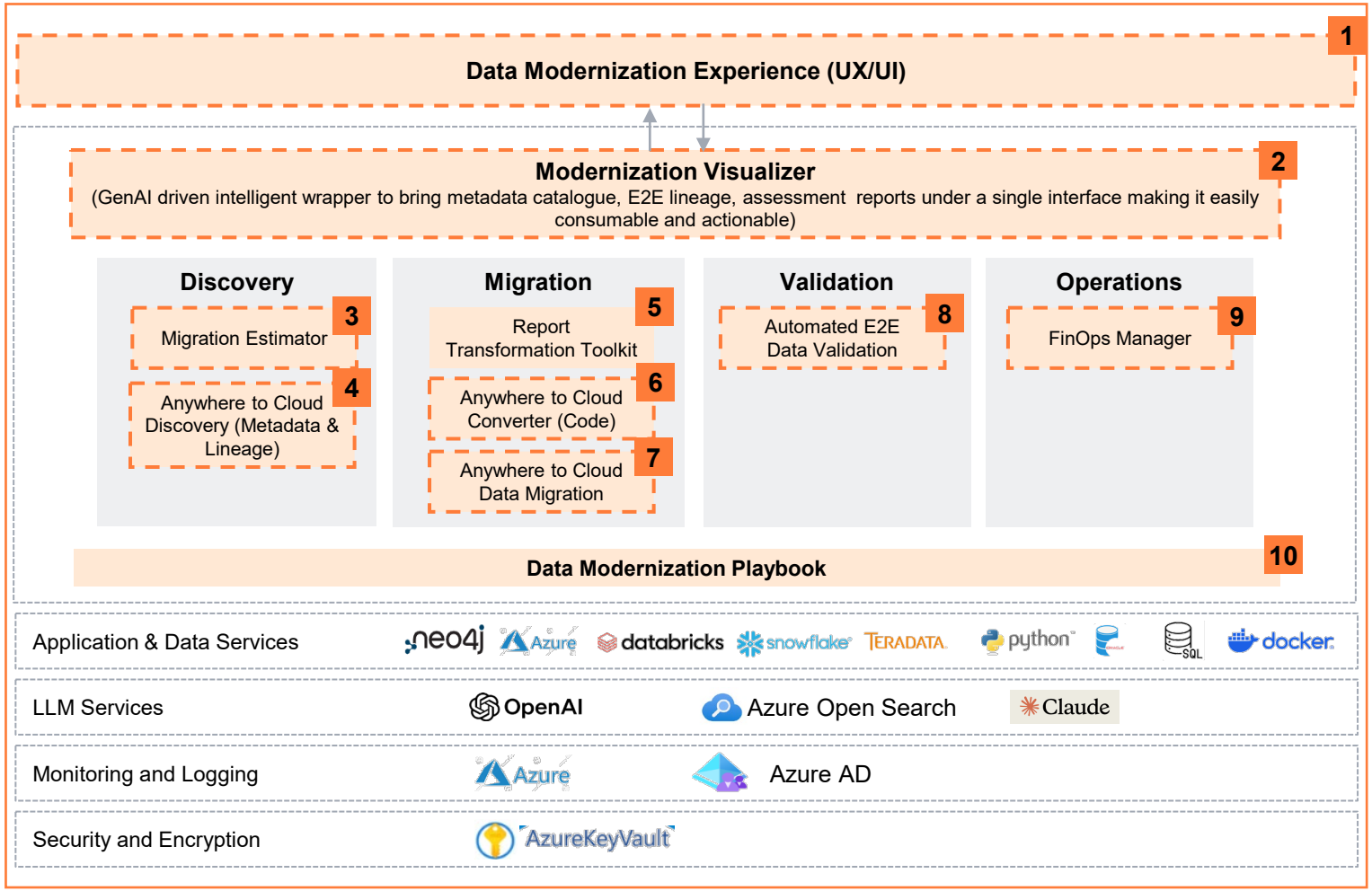
Scalable and flexible architecture

Deploy in diverse environments, including air-gapped systems, and scale horizontally to handle high-volume workloads, enabling adaptability to client specific requirements

Thorough coverage

Support a wide range of platforms, including Teradata, Oracle, SAP, Snowflake, and more, enabling modernization across various legacy and modern data ecosystems

Simplifying MassTech's journey with a proven end-to-end accelerator framework — from data discovery and migration to validation and operations*



PwC Accelerators

Accelerator Scope

1 **Data Modernization Experience (UX/UI): Inventory** - Intuitive User Interface for client demos

2 **Modernization Visualizer (RAD)**: Front-end intelligent wrapper to visualize end-to-end modernization

3 **Migration Estimator**: Automate effort estimation for migration from legacy data warehouse to modern data platform

4 **Anywhere to Cloud Discovery (Metadata & Lineage) :** Automate discovery and lineage collection from the legacy platforms

5 **Report Transformation Toolkit**: Automate report conversion from legacy to modern reporting

6 **Anywhere to Cloud Convertor (Code):** Automate code (e.g. database, ETLs, Scripts) conversion

7 **Anywhere to Cloud Data Migration**: Automate historical data migration from legacy data warehouse to modern platform

8 **Automated E2E Validation**: Automated metadata driven framework that executes the testing process for data migration and ETL

9 **FinOps Manager**: Automated process to monitor cost consumption for modern data platforms; provide insights/alerts for deviations and action items

10 **Data Modernization Playbook:** Outlines the key aspects of delivering data mod engagements incl. RFP Q&As, responses, delivery methodology, RACI, etc.,

*Illustrative Technology stack only , will be curated to MassTech during technical design sessions

Examples of PwC IDM accelerators delivering velocity, efficiency , effectiveness in delivery timelines

Modernization Stage	Accelerator	Duration
Discovery Legacy metadata, lineage, and platform dependencies are captured using automated tools to enhance visibility and reduce manual discovery efforts.	**Anywhere to Cloud Discovery (Metadata & Lineage)**: Discovery and lineage collection from the legacy platforms	60 — 40 % Reduction
	Migration Estimator: Legacy data warehouse to modern data platform effort estimator	60 — 40 % Reduction
Migration Legacy code, scripts, and data are converted into modern formats by automated tools, reducing manual effort and complexity.	**Anywhere to Cloud Convertor (Code):** Code conversion from legacy to desired target format	65 — 35 % Reduction
	Anywhere to Cloud Data Migration: Historical data migration from legacy data warehouse to modern platform	65 — 35 % Reduction
	Report Transformation Toolkit: Report conversion from legacy to modern reporting	65 — 35 % Reduction
Validation Accuracy and confidence are ensured through automated, end-to-end testing across migrated data and processes.	**Automated E2E Validation:** Metadata driven framework, executes the testing process for data migration and ETL	60 — 40 % Reduction
Operations Performance monitoring, cost tracking, and post-migration optimization are integrated for continued operational efficiency.	**FinOps Manager:** Monitor cost consumption for modern data platforms	70 — 30 % Reduction

0 — 100

Normal Duration | Duration with Automation

User Experience Enhancement

Data Modernization Playbook: Outlines the key aspects of delivering data mod engagements

Data Modernization Experience (UX/UI): Intuitive User Interface for client demos

Modernization Visualizer (RAD): Front-end intelligent wrapper to visualize end-to-end modernization

Discovery | Illuminate every data dependency in DCC with intelligent lineage mapping and AI-powered analysis

Anywhere to Cloud Discovery (Metadata/Lineage)

Automates code (e.g., database, ETLs, scripts) Metadata and Lineage extraction

Features:

- Source to target mapping for Metadata and Lineage
- Automated lineage extraction (dependency matrix)
- Impact Analysis of change propagations
- Inputs: **Oracle SQL Server, Teradata, directly upload from CSV… and more!**
- Output: **Excel extract with graph database visuals, with metadata mappings**

Expected benefits:

- Accelerate metadata analysis by providing lineage across scripts, database objects, data pipelines
- Significant reduction in effort for impact analysis (30%–50%)
- Significant reduction in effort in development and deployment planning

Migration Estimator

Technology		Object Lists	Unit Effort Hours based on complexity (Build Hours Only) --> Lift & Shift		
			Simple	Medium	Complex
Source	Teradata	Table	8	16	0
		Volatile table	8	16	0
		View	8	16	20
		Global Temporary Table	8	16	0
		Indexes	8	16	0
		Partitions	8	16	0
		Stored Procedure	8	16	24
		Function (UDF)	8	16	24
		Macro	8	16	24
		Trigger	8	16	24
		Sequence	8	16	0
		Synonym	8	16	0
		Constraints	8	0	0
		Access Rights (GRANT/REVOKE)	8	16	0
		DB link/connection configuraton (Database)	8	0	0
		BTEQ / FastLoad / MLoad(ETL)	8	16	0
		Statistics (COLLECT STATISTICS)	8	16	24
		Fallback / Backup & Recovery Plan setup	8	16	0
	ETL/ELT	Source Data Profiling	8	16	0
		Mapping Design	8	16	0
		Data Extraction	8	16	20
		Data Cleansing	8	16	0
		Workflow/Taskflow	8	16	0
		Load Scripts / Jobs	8	16	0
		Error Handling	8	16	24
		Testing & Validation	8	16	24
		Performance Tuning	8	16	24
		Reusability & Parameterization	8	16	24

Estimates the migration effort based on complexity measurements from other accelerators or based on manual inputs

Features:

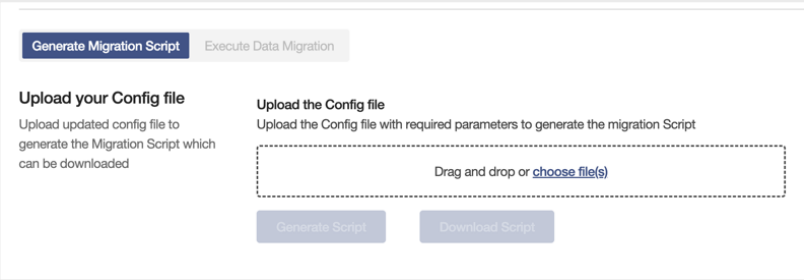
- Estimation template to estimate data mod projects in a standardized way i.e., required set of inputs, assumptions, and features to be considered
- Inputs: **Teradata**
- Outputs: **Databricks, Snowflake … and more coming soon!**

Expected benefits:

- Accelerates the effort estimation process
- Provides flexibility to try different scenarios and do sensitivity analysis and standardize and validate the estimation approach across efforts

Migration | Transform legacy code, scripts, and data into modern formats with automation that reduces manual effort and complexity

Anywhere to Cloud Data Migration


Generate Migration Script
Execute Data Migration
Upload your Config file
Upload updated config file to generate the Migration Script which can be downloaded
Upload the Config file
Upload the Config file with required parameters to generate the migration Script
Drag and drop or choose file(s)
Generate Script
Download Script

Automate historical data migration from legacy to modern data platforms.

Features:

- Generate code and migrate data from a legacy platform to a modern data platform
- Inputs: **Teradata, Qubole, Oracle, and SQL Server**
- Outputs: **Snowflake and Databricks**

Expected benefits:

- Accelerates development efforts in migrating data, with significant effort savings (30%–40% based on migration complexity)

Anywhere to Cloud Code Convertor

Automate code (e.g., database, ETLs, Scripts) conversion

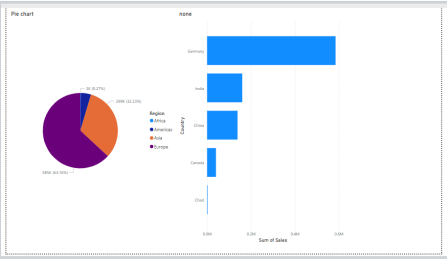
Features:

- Convert and validate code through LLMs within environment
- Generate downloadable log summaries, highlighting errors and warnings
- Inputs: Teradata, DataStage, COBOL, EasyTrieve, and Qubole
- Outputs: Databricks, PySpark, and Snowflake

Expected benefits:

- Accelerates the development effort in migrating code with significant savings in effort (30%–40% based on migration complexity)

Report Transformation Toolkit



Automate report conversion from legacy to modern reporting

Features:

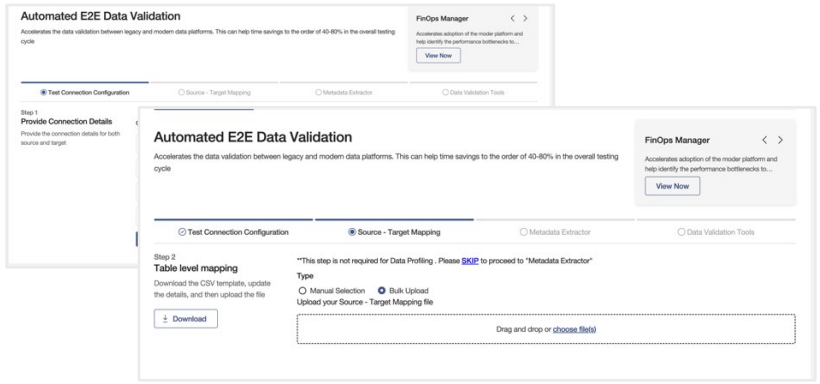
- Convert Map charts, tables, conditional formatting, drill-downs, and KPIs
- Reconstruct schemas, dimensions, hierarchies, calculated fields and measures
- Input and Output: Tableau and Power BI

Expected benefits:

- Significant savings from automating the report conversion — up to 30%–40%

Validation and Operations| Ensuring every dataset within DCC is trusted, governed, and performing at its best

Automated E2E Validation


Automated E2E Data Validation
Accelerates the data validation between legacy and modern data platforms. This can help time savings to the order of 40-80% in the overall testing cycle
FinOps Manager
Accelerates adoption of the moder platform and help identify the performance bottlenecks to...
View Now
Test Connection Configuration
Source - Target Mapping
Metadata Extractor
Data Validation Tools
Step 1
Provide Connection Details
Provide the connection details for both source and target
Step 2
Table level mapping
Download the CSV template, update the details, and then upload the file
Download
**This step is not required for Data Profiling . Please SKIP to proceed to "Metadata Extractor"
Type
Manual Selection
Bulk Upload
Upload your Source - Target Mapping file
Drag and drop or choose file(s)

Automated metadata-driven framework that executes the testing process for data migration, AI, and ETL processed

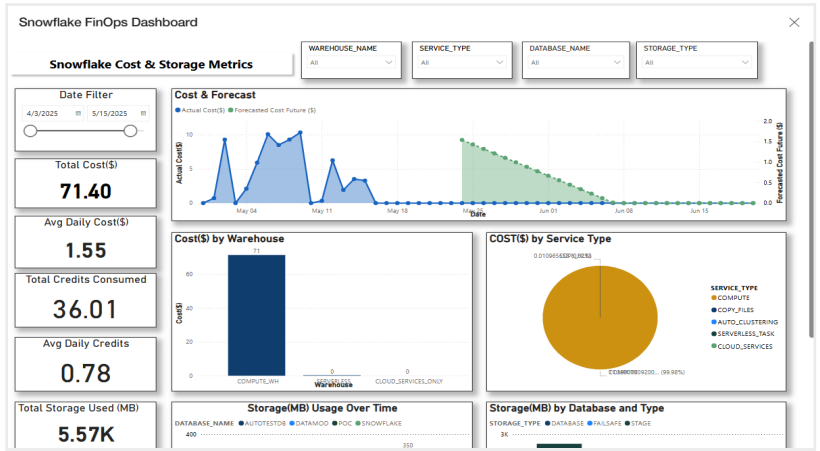
Features:

- Build schema comparator for comparing source and target schemas during as-is migration
- Perform column-level and row-by-row data comparisons for all columns after data migration
- Include custom data-profiling rules that can be applied to any source and target
- Use a React UI to capture configuration details and display results and comparisons
- Testing and monitoring of AI affords the opportunity to address risks and enable quality throughout AI lifecycle
- Inputs: **SQL Server, flat file (CSV, Parquet), Teradata, Oracle, Azure SQL DB, and Qubole**
- Outputs: **Databricks, Snowflake, and Fabric**

Expected benefits:

- Accelerates data validation between legacy and modern data platforms, enabling time savings of approximately 30–50% in the overall testing cycle
- Model stakeholders should have tests, thresholds, and controls in place that periodically monitor their performance to detect model failures or deterioration in model performance.

FinOps Manager



Automated process to measure and monitor cost consumption for modern data platforms, providing insights and alerts for deviations and items requiring attention.

Features:

- ETL Optimizer
- Snowflake FinOps dashboard
- Databricks FinOps dashboard

Expected benefits:

- Manage cloud and modernization expenses effectively
- Identify performance bottlenecks to quickly resolve issues and deliver effort savings through performance management
- Provides insights into cost trends, storage metrics, and usage breakdowns, helping teams make data-driven financial decisions

Addressing DCC Requirements | **Synthetic data generation capability**

Our Proposed Solution/Assumptions

- Implement synthetic data generation pipelines to expand access while protecting privacy.
- Ingest diverse data types (structured, unstructured, images, and time-series) through secure workflows.
- Apply de-identification, profiling, and privacy risk scoring before generation.
- Generate synthetic data using statistical models, generative AI (GANs, VAEs), and domain-specific simulators.
- Store synthetic datasets with metadata, provenance, and clear tags distinguishing them from real data.
- Maintain compliance with HIPAA, state privacy laws, and FAIR principles.

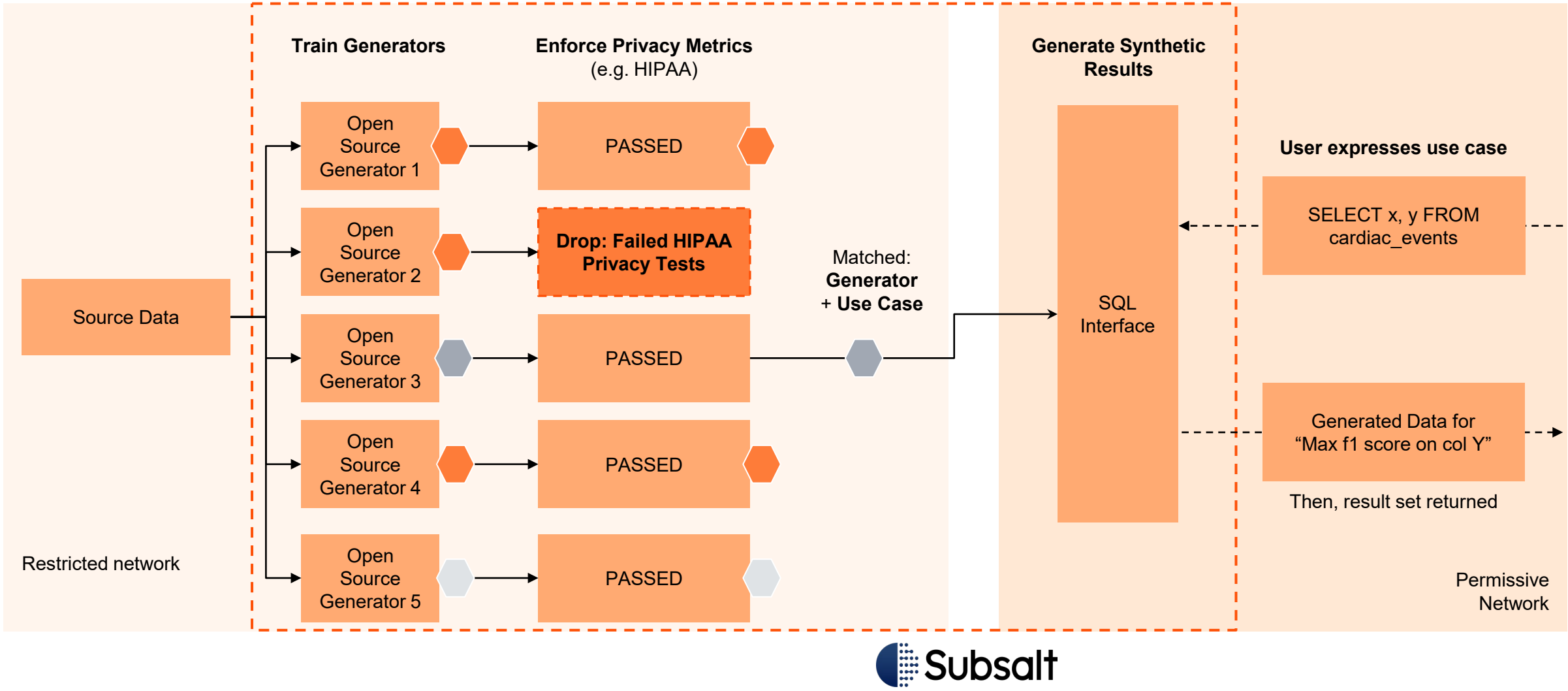
MVP Scope of Services

- Set up ingestion and preprocessing workflows for priority domains.
- Deploy open-source and commercial tools (e.g., SDV, Gretel.ai, Synthea) for synthetic dataset generation.
- Configure pipelines for fidelity vs. privacy tradeoff management.
- Establish metadata linkage between original and synthetic datasets for traceability.
- Stand up containerized execution environments for reproducibility.
- Define governance model for synthetic dataset creation, sharing, and auditing.

PwC Tools and Accelerators

- Custom synthetic data tools and code bases, integrated with Subsalt technology stack

Subsalt Automates the **Model + Use Case** Matching and Enforces Regulatory Privacy Thresholds



Subsalt Automates the **Model + Use Case** Matching and Enforces Regulatory Privacy Thresholds

Data ingestion & Preprocessing

- Supports **structured, tabular data** for both cross-sectional and time-series generation
- Future plans to support unstructured data (e.g., free text such as doctor's notes)
- System scans connected data sources and provides **automated suitability assessment**
- **HIPAA-certified, de-identified synthetic data** generated automatically (no preprocessing required)
- Maintains **Expert Determination** from Datavant

Synthetic Data Generation

- Provides a **compliance wrapper** around open-source and commercial models (GANs, VAEs, etc.)
- For anonymization use cases, **privacy thresholds are fixed to regulatory standards** (e.g., HIPAA)
- On-demand, **use-case specific synthetic data generation**
- Library of pre-trained generators selected in real time for each query
- Removes user need to configure parameters, reducing privacy risk

Auditability &Traceability

Full audit logs for:

- Data sources used
- System configurations
- Generators employed
- Timing of synthetic data creation
- Query patterns and similarity metrics

Privacy and utility metrics tracked for each generator

Storage & Metadata

- Synthetic data delivered **on-demand** into analytical environments
- Reduces need for duplicative storage of synthetic vs. real data
- Metadata (privacy metrics, utility metrics, usage patterns) stored in-system
- **Role based access control** for metadata visibility

Compute Infrastructure

- Uses common **cloud GPU/CPU resources**
- Fully **containerized application**
- Deployable into **customer environments** including **air-gapped installations**

Privacy & Compliance

- Maintains **HIPAA Expert Determination** across system and processes
- All data programmatically meets HIPAA de-identification standards
- **Optional Differential Privacy** (configurable per dataset)
- Only HIPAA-compliant synthetic data retrievable

Access Control & Governance

- Fine-grained user roles: from **Read-only** to **Administrator**
- Admins can configure system, organizations and users
- Enables broader dataset sharing under **controlled governance**

Subsalt Platform Overvies

- **Orchestration system** for compliant synthetic data workflows
- Produces only **HIPAA-compliant datasets** via Expert Determination
- Generative method-agnostic: tied to **outcomes and guardrails, not models**

Subsalt Automates the **Model + Use Case** Matching and Enforces Regulatory Privacy Thresholds

Automated End-to-End Pipeline

- Secure ingestion of source data & metadata
- Isolated compute provisioning for each job
- Enforcement of privacy policies
- Automated model selection & parametrization
- Post-generation **utility + privacy risk evaluations**
- Issuance of **privacy & quality reports**

Outputs & Delivery

- Exposed primarily through **PostgreSQL interface** (query synthetic data via SQL)
- Optional **flat-file CSV exports**
- All requests, model selections and reports **recorded in audit logs**

Compliance & Guardrails

- Anonymization guarantees **programmatically enforced**
- Downstream contractual restrictions applied at delivery
- Ships with curated open source + commercial generators
- **NDA-protected appendix** of deployed generators available
- Flexible to adopt **new generative methods** over time

Addressing DCC Requirements | **AI fairness and bias capability | Governance policies and ethical guidelines**

Our Proposed Solution/Assumptions

- Establish fairness and bias guardrails for data within the Commons, focusing on bias in input datasets rather than in models.
- Leverage PwC's Responsible AI Framework to define fairness dimensions, bias detection approaches, and acceptable data-use principles.
- Embed observability hooks and data quality checks now, enabling future AI services to extend seamlessly to bias and monitoring at the model level.
- Assume extensibility is critical — the platform must be designed so that future AI/ML use cases can integrate with fairness testing with minimal re-architecture.

MVP Scope of Services

- Define AI governance design, including bias and fairness requirements for datasets ingested into the Commons (e.g., demographic representativeness, protected attributes).
- Integrate baseline bias scanning and profiling controls at the dataset level using open-source platforms.
- Provide a template dataset of "fairness and bias cards" as part of the metadata catalog to document potential bias sources and data limitations.
- Develop requirements for observability integration (e.g., Datadog) so that future AI models can be monitored for bias, drift, and explainability.

PwC Tools and Accelerators

- PwC Responsible AI Framework and AI Policies – Enables ethical, transparent, and accountable AI governance across the full lifecycle.
- PwC AI Risk Model Cards – Enable traceable, explainable, and auditable AI model documentation and oversight.
- JBRs with Observability Platforms (e.g., Datadog) – Enable real-time monitoring, performance tracking, and model drift detection.
- PwC AI Bias Toolkit – Enables detection, measurement, and mitigation of bias in data and AI models.
- PwC Model Testing/AI Evaluation Toolkit – Enables robust validation and performance assurance of AI models before deployment.

Addressing DCC Requirements | **Governance requirements**

Our Proposed Solution/Assumptions

- Establish a DCC governance framework spanning legal/regulatory, ethical, and technical controls, inclusive of high-risk use cases.
- Define roles, processes, and escalation paths (data stewarding, fairness review, approvals, exceptions) mapped to 3-line defense.
- Specify data-sharing models and access tiers with clear provider–user responsibilities and DaaS terms, including federated training/validation options.
- Implement consent governance (tokenized, revocable, time-bound) and acceptable-use rules; integrate nondiscrimination and bias-risk safeguards.

MVP Scope of Services

- Define governance platform requirements, including for AI.
- Consent management blueprint with revocation and purge workflows; linkage to dataset and derivative tracking.
- Charter the Data Access and Use Committee for oversight.
- Translate policy to platform requirements: audit logs, lineage, approvals, encryption, RBAC/MFA, and public-facing transparency summaries where appropriate.
- Monitoring & reporting with KPIs, dashboarding for reviews/approvals, fairness summaries, and public transparency reports.
- Document acceptable use framework for data and AI.

PwC Tools and Accelerators

- PwC Compliance Management System (CMS) Framework
- Policy-as-Code patterns
- PwC Sample AI Roles and Responsibilities and RACIs
- PwC Responsible AI Principles & Framework
- PwC AI Bias, Model Testing, & Evaluation Toolkit
- PwC Draft Policies & RACI Templates
- JBRs with Observability Platforms (e.g., Datadog)
- PwC Transparency & Assessment Toolkit

Enabling the responsible adoption of AI requires an enterprise risk and governance framework that can be integrated into the Digital Data Commons

Core elements of a Responsible AI Framework

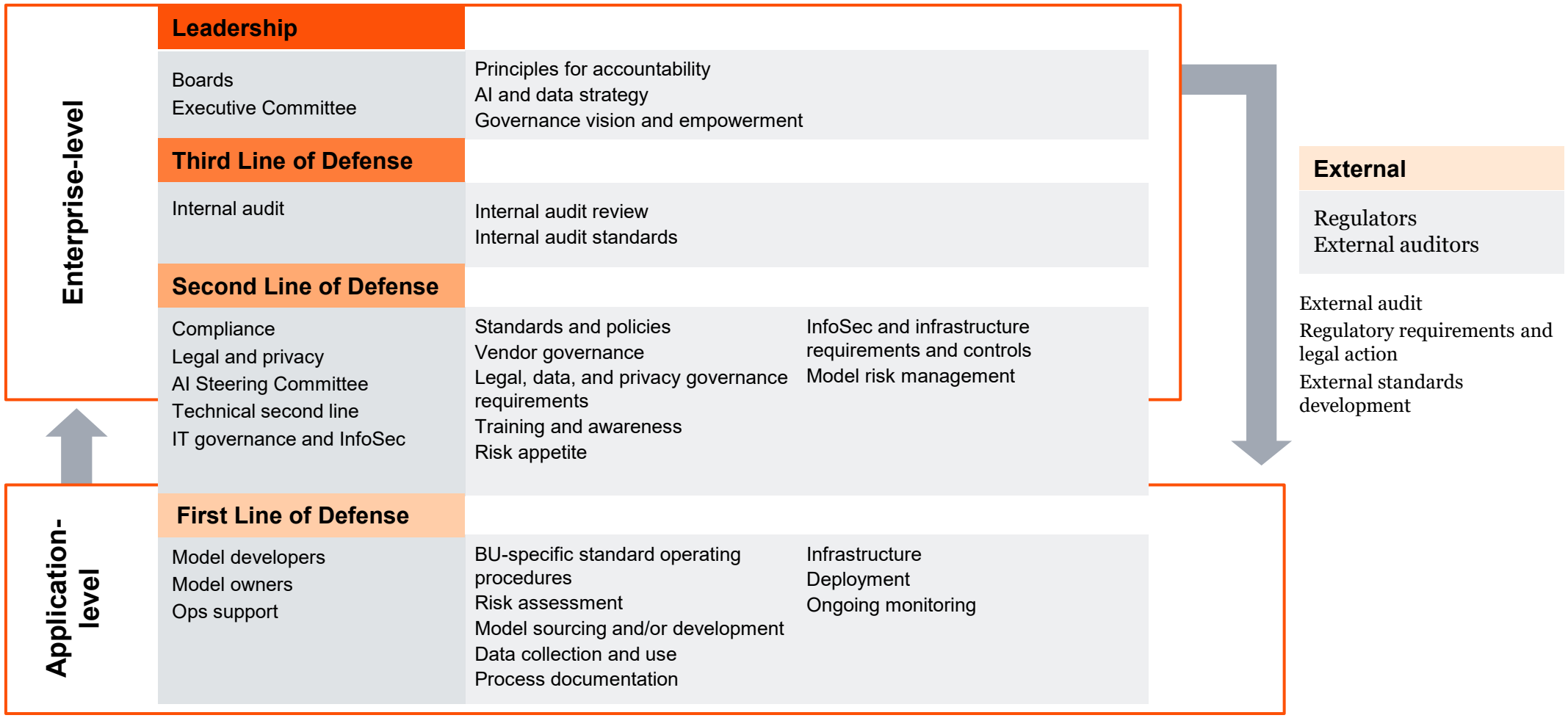
Foundational Capabilities	Operating Model and Governance	Application Life Cycle
AI Principles and Strategy	Operating Model: Roles and Responsibilities	AI Development and Deployment Standards
AI Use Case Inventory	Governance Committee and Escalations	AI Testing and Monitoring
AI Risk Taxonomy	AI Risk and Control Matrix	Risk Mitigation Tracking and Reporting
AI Risk Intake and Tiering	Training and Communication	

Policies and Procedures Across Risk Domains (e.g., cyber, privacy, legal, model risk)

Developing responsible AI and governance is not just one person's job, and different models should be considered for the Data Access and Use Committees

Diverse stakeholders across all parts of the organization should collaborate to apply Responsible AI practices consistently and effectively.

The Three Lines of Defense

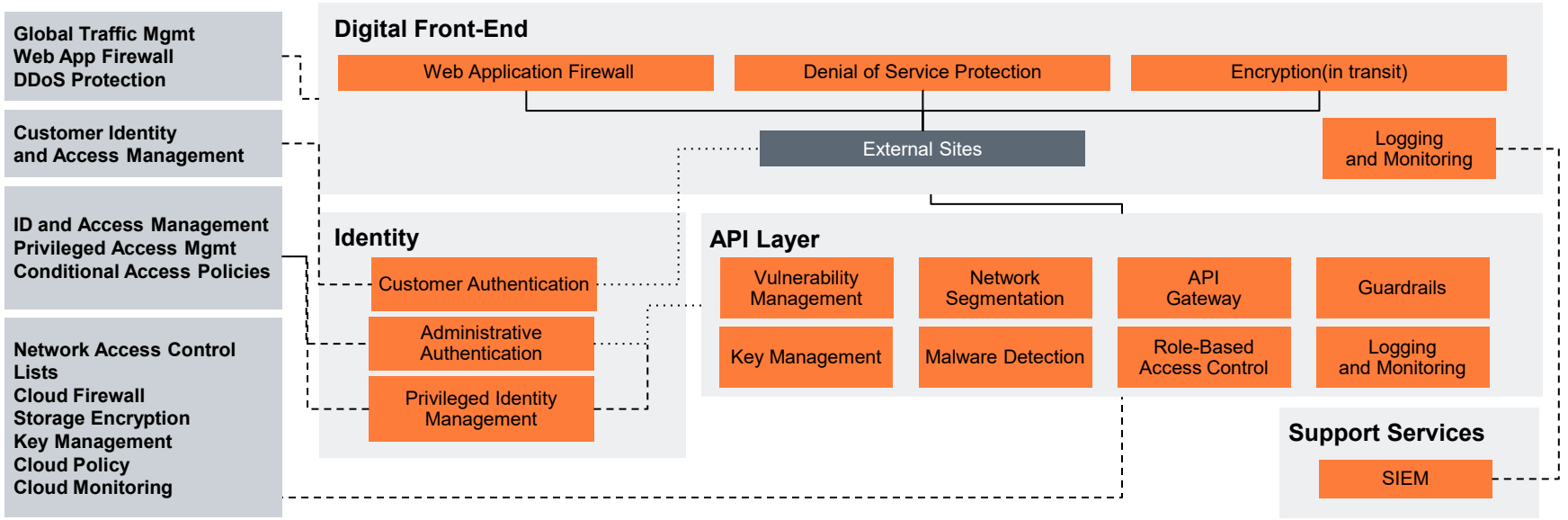


Secured and trusted through Security, **Privacy, and Compliance**

Embed Security-by-design and privacy-by-design, trusted by some of the world's largest institutions, to safeguard data across all processes. Maintain Privacy and Compliance through encryption, access control, incident response, regular audits, and adherence to responsible AI principles while adhering to Security by design principles.

Illustrative "defense in depth" approach to embed security in the Commons data platform

Services Required:



Secure by design throughout the life cycle of development

Threat modelling: We'll analyze the target environment and, using methodologies like STRIDE, define threats against the target architecture.

Capability maps: Using the threat model and your reqs, we'll identify and define security capabilities to mitigate those threats.

Standards and Engineering Principles: The insights will inform design and engineering principles for secure implementation.

Things to note:

- Implement advanced security technologies (e.g., encryption, AI-driven threat detection) to safeguard **data at rest, in transit, and during processing**. Include real-time monitoring and response capabilities.
- Set up an **incident response framework** that quickly identifies, mitigates, and learns from breaches, establishing operational resilience.
- Implement **Responsible AI principles** to secure traceability, accountability, transparency, and explainability for future uses of AI.
- Establish a clear **data classification system** to differentiate between levels of sensitivity, making sure that the right security measures are applied to each type of data.
- Enable **platform-level observability** through instrumentation of logs and critical events across the common data platform using DataDog.

Addressing DCC Requirements | **Security and access controls**

Our Proposed Solution/Assumptions

- Portal UI protection with WAF and securing integrations with segmentation and firewalls
- API gateways for protecting API integrations
- MFA for authentication and Single Sign-On via OIDC/OAuth2/SAML with IdP-initiated and SP-initiated flows; SCIM for life cycle
- Encryption at rest with cloud KMS/HSM
- Secure engineering principles with risk-based stage gates, code scans, and scheduled penetration tests
- Platform resiliency across the solution stack
- Defense in depth by implementing industry-leading security practices across Commons data platform

MVP Scope of Services

- MFA
- SSO
- RBAC-based access to data sets
- Integration of critical logs into enterprise SIEM for overall platform observability
- Encryption of data at rest and in transit
- Security by design principles embedded in the development life cycle
- Baseline privacy by design principles established
- Baseline platform resiliency

PwC Tools and Accelerators

- PwC Responsible AI Framework
- Template Policies

What is observability in data and AI systems?

It is the ability to explain AI systems by turning raw signals into evidence of why outcomes happen; making complex models, agent workflows, and emergent behaviors transparent, accountable, and controllable in real time.

Key dimensions	Understanding AI observability	Examples
System Behavior	What is the AI doing, is it biased, and why is it making these decisions?	• Prompt/Response logs • Model decision traces • Agent workflows and chaining behavior
System Health	Is the AI performing reliably and at the expected quality?	• Latency and throughput • Error rates • Resource utilization (CPU, GPU, memory)
User Access	Who is using the system, and how are they interacting with it?	• Authentication and access logs • User roles and permissions • Session and usage activity
Data Flow	What data is being used, where is it going, and what is produced?	• Input/Output lineage • Source access logs • Volume and frequency of data movement

Observability reveals what's happening inside AI systems — how they decide, perform, and use data.

By collecting and connecting this evidence, organizations can spot risks early, reduce hallucinations, build trust, and utilize AI safely.

Control implementation and confidence with observability

We design observability to make AI systems transparent by converting raw signals into auditable controls that prevent issues, detect risks, demonstrate compliance, and strengthen governance.

Model/Agent Governance

Record versions, routing, drift, and bias; enforce guardrails and quarantines.

Data Governance

Track lineage, completeness, accuracy, redaction, and error analysis to expose sources of data issues.

Human-in-the-loop

Provide humans with context to validate predictions, correct errors, and make final decisions.

Access and Security

Enforce role-based access to data, infrastructure, and AI agents by limiting permissions and monitoring activity.

Auditability

Maintain immutable logs as continuous evidence to enable the ethical and lawful use of AI systems.

Incident Management

Tie alerts to service-level objective burn rates and generate evidence packs to accelerate root cause analysis.

Change Management

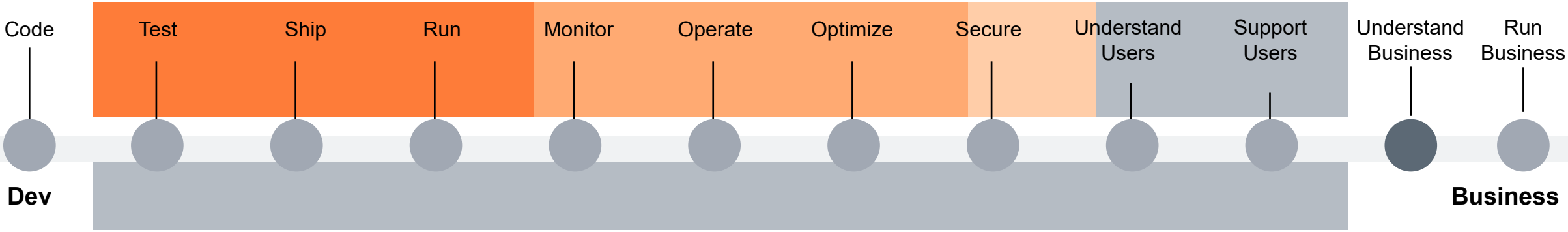
Capture and correlate each deployment, prompt, model, tool, or flag change with outcomes.

Platforms like Datadog give leaders real-time visibility to spot issues, pinpoint causes, assess impacts, and act quickly. They surface emerging problems as they arise, enabling proactive prevention before escalation

We have a JBR with Datadog, which includes a trial license to be provided once ready to implement.

Delivering continuous monitoring and assurance through PwC's partnership with Datadog

Illustrative "Observability" capabilities to embed security and resiliency in the Commons data platform



Software Delivery

- CI Visibility
- Test Optimization
- Continuous Testing

Monitor and Operate

- Infra Monitoring
- Network Monitoring
- APM
- Synthetics
- Log Mgmt
- Universal Service Monitoring
- Observability Pipelines
- LLM Observability

Enhance

- Continuous Profiler
- Cloud Cost Mgmt
- Database Monitoring
- Data Streams Monitoring
- Data Jobs Monitoring
- K8s Autoscaling

Secure

- Cloud Security Mgmt
- Code Security
- Cloud SIEM
- Sensitive Data Scanner
- Workload Protection
- App and API Protection

Analyze

- RUM
- Heatmap/Clickmap/Scrollmap
- Product Analytics
- Mobile App Testing
- Session Replay

Cloud Service Management

- Incident Management
- On-Call
- Workflow Automation
- App Builder
- Software Catalog
- Resource Catalog
- Case Management
- Event Management

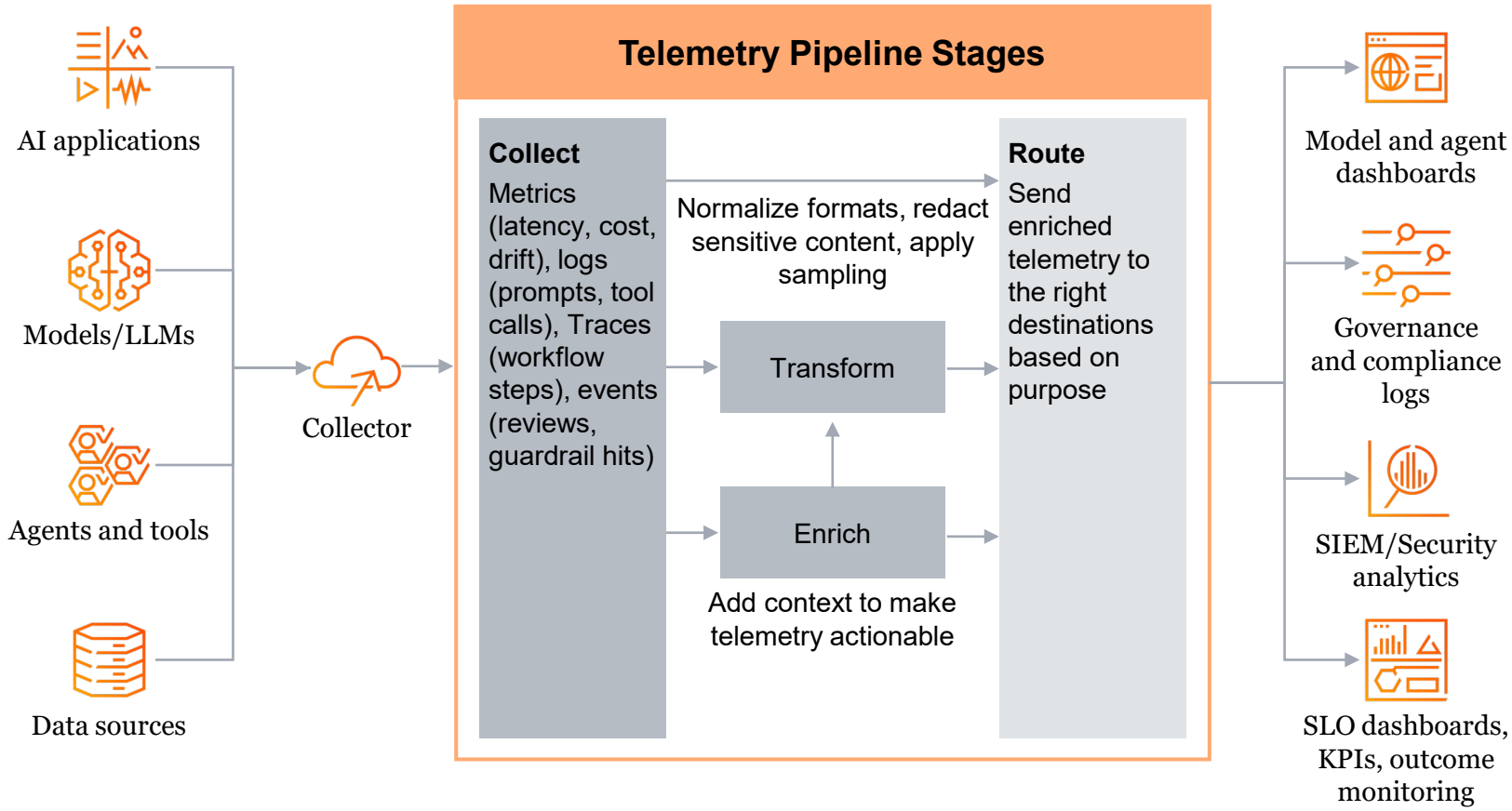
A long-term responsible and ethical AI platform requires a telemetry pipeline

AI and data telemetry captures metrics, logs, traces, and events from models, agents, and data sources. Which is then used in a pipeline that collects, enriches, and routes this data to dashboards, compliance logs, and security tools enabling reliable, auditable AI observability.

Sources

Telemetry begins with AI applications, models/LLMs, agents and tools, and data sources. Each of these generate signals:

User interactions, sessions, prompts, token usage, response outcomes, reasoning steps, tool calls, retries, retrieval queries, lineage, and freshness checks



Destinations

Telemetry flows out to multiple endpoints for analysis and decision making:

Monitor performance, cost, drift, maintain immutable audit trails, policy checks, regulatory evidence, detect anomalies, prompt injection, track success rates, and other business-aligned objectives

Appendix

Biographies

Rima Safari
Advisory Partner
+1 (201) 710-0813
rima.safari@pwc.com

What I will do:

The buck stops with me, I will serve as your value realization officer, working very closely with MassTech, the core team, and specialists on the delivery of your AI Hub vision.

Experience:

- Rima is currently leading one of the most complex transformations across the group insurance industry for a Fortune 200 Insurer – Unum. This is a $30 million initiative this year, #1 priority investment for their CEO and expected to continue with over $40 million investment for the next two years. Rima is the face of this program, leading a team of over 20 PwC resources at its peak, across FINEOS implementation, architecture, customer experience, PMO, etc.
- She worked with a leading L&A client to consolidate their global IT environment, by migrating operations from 14 countries to three facilities in Asia Pacific, Europe, and Latin America . The migration process was a complex undertaking where she led the planning and scoping track, coordinated with multiple project managers and infrastructure teams across the globe to come up with the best migration approach and ensure timely execution.

Oneil Rego
Advisory Partner
+1 (267) 987-4012
oneil.rego@pwc.com

What I will do:

I will serve as your senior relationship partner, coordinating closely with the core team and AI/Data specialist to make sure PwC is delivering high-quality services, bringing the depth and breadth of resources, and delivering value to you.

Experience:

- Led CIOs in developing organizational transformation strategies and execution to align with key business priorities over the short, mid, and long term
- Led a large analytics transformation program to help implement an analytics platform and backbone for a regional payer, including interoperability and informatica, for member mastering as well as provider data exchange
- He has more than 28 years of experience advising executives at pharmaceutical, biotechnology, and medical devices companies, focused on business operations integration and divestiture, sales and marketing, intelligence strategies, and company launch and growth.
- He has a track record of delivering complex initiatives focused business value creation, and consensus building through the development and execution of business visioning workshops.

Biographies

Ravi Kalakota
Advisory Partner
+1 (646) 887-7012
ravi.s.kalakota@pwc.com

What I will do:

I will serve as your lead delivery partner, working very closely with MassTech, the core team, and specialists on the delivery on DCC MVP and beyond

Experience:

- Ravi has over 28+ of experience in data operations and transformation consulting. Ravi focuses on AI, Generative AI, hybrid cloud platforms, data quality and governance, and data engineering operations. Ravi has experience leading multiple modernization efforts at national and regional plans to develop and enhance data and analytics capabilities.
- Ravi's focus is on Digital, data and AI transformations that are top-of-mind for healthcare payers and providers everywhere to address Revenue, Quality, Affordability, Member Engagement, and Care Delivery. Almost every firm has done, is doing, or plans to transformation these processes and functions.
- He is into turning strategies into practical execution and capability building. What are the critical elements it takes to rewire an organization through making fundamental changes to talent, operating model, and technology and data capabilities.
- Global CIO and CTO for a Fortune 500 company and developed IT strategies, Cloud Strategies and operating models for numerous companies.

Illana Blumenfeld
Responsible AI, Assurance Partner
+1 (847) 912-0250
ilana.a.golbin@pwc.com

What I will do:

I will serve as the lead AI Governance and Responsible AI workstream, that will draft your AI Governance policy, design, and controls for the DDC.

Experience:

- Ilana is a principal with PwC's Digital Assurance and Transparency, Tech, and Data team, where she co-leads AI. She is part of the of the PwC AI leadership and serves as PwC's Responsible AI Research lead.
- She is a Certified Ethical Emerging Technologist, is listed as one of 100 "Brilliant Women in AI Ethics" in 2020 and has been recognized in Forbes as one of 15 leaders advancing Ethical AI. She is a practicing data scientist with more than 15 years of experience in building and assessing AI/ML for clients and for PwC.
- Since 2018, she has led PwC's efforts globally in the research and development of cutting-edge approaches to build and deploy Responsible AI. Her research is published in academic literature. She is a frequent speaker and author on Responsible AI topics and works across all industries, and has advised governments, policy makers, and regulators on AI.
- Contributor to NIST AI RMF and NIST Generative AI, WEF AI (including global AI governance), IAPP, and IAF working groups, among others.

Biographies placeholders

Sachin Bajpai
Advisory Managing Director
+1 (949) 383-7812
sachin.bajpai@pwc.com

What I will do:

I will serve as your Data Architect, working very closely with the MassTech Technical Architecture team, the core team, and specialists on the delivery of a strong foundation for your DCC services

Experience:

- Sachin is a seasoned data and analytics consultant and technologist at heart.
- He is striving to simplify healthcare IT and improve outcomes through effective use of data and analytics.
- Leading the engagement to build a Data Science workbench enabling Payer Analytics use cases.
- Machine learning models for Readmission, Potentially Preventable Admissions, Long Term Services and Support (LTSS) Transition to expensive facilities, Quality Analytics, Appeals and Grievances.

Praveen Krishna
Advisory Director
+1 (314) 363-6997
praveen.krishna@pwc.com

What I will do:

I will serve as your Engagement delivery lead , coordinating closely with the core team and specialists to make sure PwC is delivering high-quality outputs, bringing the depth and breadth of resources, and delivering value to you.

Experience:

- Developed and executed a innovative design for a smart city around healthcare interoperability, business service, cybersecurity and privacy in a intense and challenging environment to reimagine healthcare
- Developed and delivered a consumer health transformation platform with multiple ecosystem partners, including overseeing managed services and operations of the platform; integrated remote patient monitoring data with clinical and behavioral data; data de-identification; and aggregation for analytics and outcomes measurement
- Led innovative approach to evidence generation leveraging real world data from hospital systems in the context of a pharma study. Enabled the platform, provider - pharma collaboration and secure containers to perform the analysis while securing the interests of both parties.
- Led a multi-year Pharma R&D transformation to deploy a translational research insights generation platform bringing in genomics and study data, creating data engineering pipeline, data orchestrations, managed services around data operations, cloud operations, security operations and solution management

Biographies

Senthilkumar Jayachandran
Advisory Director
+1 (201) 616-9276
senthilkumar.jayachandran@pwc.com

What I will do:

I will serve as your Data Platform SME, bringing outside-in perspectives of leveraging AI in driving your platform build and growth .

Experience:

- Senthilkumar has over 15+ years of consulting experience in design, delivery, and implementation of data solutions to help resolve complex business problems.
- Demonstrated ability to think outside the normal solutions box and have developed innovative solutions Responsible for developing business solutions to meet and exceed the customer needs.
- Hands-on experience at multiple roles – Technical owner, Solution Architect, Technical Lead, Data Analyst, Programmer/Support & Project Management – working with Cross-Functional, Cross-Geography & Multicultural teams of 5 -12 members across North America, South America & Asia regions.

Amol Chaudhari
Advisory Director
+1 (973) 437-6261
amol.chaudhari@pwc.com

What I will do:

I will serve as your day-to-day lead security director working across the Data commons platform to ensure security by design in every aspect of the platform.

Experience:

- Created a threat-based Gen AI Security framework for an energy firm's Gen AI initiative, incorporating MITRE Atlas, NIST 800-53 frameworks, and Responsible AI principles.
- Designed a dynamic multi-hyperscaler cloud security framework for a large cloud initiative that encompassed over 80 business solutions, including AI/ML applications. This framework facilitated one-click Infrastructure as a Code deployment of security controls, ensuring compliance with NIST 800-53, HIPAA, and PCI-DSS standards.
- Directed a security program for one of the top five global banks, comprising more than 15 projects aimed at securing and strengthening the enterprise security posture across various OSI layers. He coordinated efforts to consolidate individual services and tools, thereby enhancing the overall global security stance.
- Conducted multiple security assessments for cloud and on-premises environments, including Zero Trust architecture evaluations for clients in the Energy, Retail, Education, and Consulting Services sectors.

Biographies

Martin Coffman
Assurance Director
+1 (469) 441-7542
martin.d.x.coffman@pwc.com

What I will do:

I will serve as the director leading the day-to-day work of AI Governance and Responsible AI under Ilana.

Experience:

- Martin is a strategic leader in AI and data analytics with 13 years of experience delivering high-impact digital transformation at the intersection of people, process, and technology.
- He has a reputation for helping clients integrate AI, governance, security, and controls seamlessly into their operational workflows, making sure that outcomes are impactful, transparent, and ethically aligned.
- He has spearheaded AI adoption across PwC's business operations, overseeing sizeable investments that redefined workflows and accelerated enterprise-wide efficiency.
- His expertise spans automotive, retail, finance, healthcare, and technology industries, where he has designed AI-powered solutions, ERP strategies, and automation frameworks that have generated millions in organic growth.

Spoorthi Purumula
Advisory Senior Manager
+1 (623) 850-0783
purumala.spoorthi@pwc.com

What I will do:

I will serve as your day-to-day Data Platform lead, provide direction and oversee delivery of platform features and capabilities

Experience:

- Established over 30 data domains to effectively manage and govern data in the enterprise platform.
- Onboarded stakeholders to their ownership roles for domains - Business Data Owner, System Owner, Process Owner.
- Developed key data governance policies - Data Ownership, Data Access, Data Quality, Data Archival & Retention.
- Reconciled business requirements for harmonized complaints process and translated them into technical requirements for TrackWise implementation.

Biographies

Sean Allen
Advisory Manager
+1 (617) 388-3017
sean.r.allen@pwc.com

What I will do:

Based in Boston, I will serve as the day-to-day delivery manager for the Data Marketplace and AI Workbench design, build, and release.

Experience:

- Delivered a custom-built enterprise data marketplace on AWS for a global organization post-spinoff
- Directed the migration of reporting systems to a unified cloud platform post-merger, enabling operational scalability
- Built a roadmap for transitioning from a legacy enterprise data warehouse to a cloud-based modern data platform with AI capabilities
- Led Agile-based analytics delivery for a new enterprise analytics function, prioritizing business demand and actionable insights
- Certified SAFe® 6 Lean Portfolio Manager, Scaled Agile
- Master of Science in Quality Management, The George Washington University
- Master of Science in Informatics, Johns Hopkins University (pending capstone)

Aamir Polli
Advisory Senior Associate
+1 (872) 258-2120
aamir.polli@pwc.com

What I will do:

I will support the development of the DCC solution and help manage the build and implementation, ensuring alignment between business requirements and technical delivery.

Experience:

- Aamir is a Senior Associate in PwC's CEDA practice, specializing in data strategy, architecture, and analytics delivery.
- Proven record leading enterprise-level data programs across healthcare, real estate, and financial services sectors.
- Expertise in data modeling, data architecture, and governance — driving adoption of cloud and advanced analytics platforms such as Azure, Databricks, and Snowflake.
- Strong background in bridging business and technical perspectives, enabling the delivery of effective, future-ready data solutions.
- Based in Chicago, IL, Aamir holds a B.S. in Computer Science and an MBA from DePaul University.

Biographies

Sandipan Kumar Dey
Advisory Director
+91 8013-690857
sandipan.k.dey@pwc.com

What I will do:

My role is to oversee the team in India that will deliver on your Data Marketplace, following quality and due process to deliver a high performing solution.

Experience:

- AWS certified solutions architect with over 14 years of experience in IT industry & hands-on project experience in Big data Hadoop technology stack like SPARK with Scala, HDFS, HBase, Hive, Pig, Sqoop, Oozie , AWS technologies like Redshift, S3, Kinesis Firehose and Azure technologies like Azure Databricks, Azure Data Factory & Azure DataFlow**.**
- Expertise in Oracle SQL, PL/SQL, Syncsort DMExpress ETL tool and Linux Bash shell scripting
- Skilled in all phases of the software development lifecycle; expert in translating business requirements into technical solutions; and fanatical about quality, usability, security and scalability.
- Experience in working with AWS components like Macie, Lambda, S3, Redshift, Kinesis Firehose, Glue, SNS, SES, Cloudformation, CloudWatch, Secret Manager, RDS, DynamoDB, Step Function.

Srikanth Kallepalli
Advisory Manager
+1 (918) 640-6553
srikant.kallepalli@pwc.com

What I will do:

I will serve as the PwC Tech lead , based in India, overseeing the quality and output of data platform and cloud engineers and ensure they are aligned to your goals and objectives

Experience:

- Srikant is an Experienced Associate with PwC's Mgmt. Consulting HIA practice. He has over 4 years of professional experience in US healthcare industry, with focus on US health payer.
- Srikant currently holds a Professional, Academy for Healthcare Mgmt. (PAHM) designation.
- He completed his Post Graduate Diploma in Management-Finance from Institute of Management Technology (IMT) Ghaziabad, India and Bachelor of Tech. in Electronics and Communication Eng. from SRM University, Chennai, India.
- Data & Analytics strategy to establish an Innovation Pod for a large health payer to engage with evolving vendors; also, research on the evolving AA & ML vendors across areas like Care Mgmt., Claims Mgmt. Ops Excellence etc.

Thank you

pwc.com

© 2025 PwC. All rights reserved. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.

REFERENCE FORM

Full Company Name: University of Texas System (previous) | Apricity Health (current)
Address: 3262 Westheimer Road, Suite 762, Houston, Texas, 77098
Name/Title of Contact Person: **Dr. Lynda Chin, MD** Department chair and professor of genomic medicine at UT system Co-founder, chief executive officer, chairperson of board at Apricity health
Phone Number/Email of Contact Person: PwC will facilitate conversation upon request
Email Address: PwC will facilitate conversation upon request

Full Company Name: Eli Lilly and Company
Address: 1094, Lilly Corporate Ctr, Indianapolis, Indiana, 46285-0001
Name/Title of Contact Person: Ken Supanick / Head of Enterprise Data, Bob Taylor / Enterprise Data Marketplace Lead
Phone Number/Email of Contact Person: PwC will facilitate conversation upon request
Email Address: PwC will facilitate conversation upon request

Full Company Name: Auto Desk
Address: 1 Market Street, San Francisco, CA 94105
Name/Title of Contact Person: Patricia France, Senior Director of Trust Governance, Risk & Compliance
Phone Number/Email of Contact Person: PwC will facilitate conversation upon request

Email Address: PwC will facilitate conversation upon request

Massachusetts Technology Collaborative

Grant Agreement
Standard Budget and Invoice Template
Invoiced Costs

Line Number	Budgeting															
2	Applicant Information															
3	Applicant:			Procurement No.:												
4	Budget Period:			Title of Proposed Project:												
5	Address:															
6																
7																
8				MTC Funding Requested:												
9				Total Project Cost:												
10	Cost Elements														Cumulative	
11															Period: MM/DD/YYYY-MM/DD/YYYY	
12															Expended	Remaining
13	I. Direct Labor															
14	*name/title*	*HRS*	*Rate*		*HRS*		*HRS*		*HRS*		*HRS*					
15				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
16				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
17				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
18				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
19				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
20				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
21	Total Direct Labor			$0	0	$ -	0	$ -	0	$ -	0	$ -	0	$ -	$ -	$ -
22																
23	II. Direct Labor Fringe Cost	Rate (%):		$0		$ -		$ -		$ -				$ -	$ -	$ -
24																
25	II. Subcontractors/Consultants/Contracted Staffing															
26	*Firm name/title or role/type of services to be provided*	*HRS*	*Rate*		*HRS*		*HRS*		*HRS*		*HRS*					
27				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
28				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
29				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
30				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
31				$0	5	$ -		$ -		$ -		$ -		$ -	$ -	$ -
32				$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
33	Total Subcontractors/Consultants			$0	5	$ -	0	$ -	0	$ -	0	$ -	0	$ -	$ -	$ -
34																
35	III. Direct Materials (provide description and costs basis)															
36	Curriculum Development & Training			$0											$0	$0
37				$0											$0	$0
38				$0											$0	$0
39				$0											$0	$0
40				$0											$0	$0
41				$0											$0	$0
42				$0											$0	$0
43	Total Direct Materials			$0		$0		$0		$0		$0		$0	$0	$0
44																
45	IV. Travel														$0	$0
46																
47	V. Other Direct Costs *(list type and description such as venue rental costs and cost basis)*															
48				$0											$ -	$ -
49				$0											$ -	$ -
50				$0											$ -	$ -
51				$0											$ -	$ -
52				$0											$ -	$ -
53	Total Other Direct Costs			$0		$ -		$ -		$ -		$ -		$ -	$ -	$ -
54																
55	VI. General & Administrative Expense/Overhead	Rate (%):		$0		$ -		$0		$0		$0			$ -	$0
56																
57	**Total PROJECT COSTS**			**$0**		**$0**		**$0**		**$0**		**$0**			**$0**	**$0**

LEGEND:

Formula cell
Data entry cell

This section is for post-contract invoicing only. Do not complete with proposal.

Please refer to Section 6 - Project team and Fees in the main proposal response document - pages 41 through 43